Frontiers of Information Technology & Electronic Engineering
www.jzus.zju.edu.cn; engineering.cae.cn; www.springerlink.com
ISSN 2095-9184 (print); ISSN 2095-9230 (online)
E-mail: jzus@zju.edu.cn

Distributed game strategy for unmanned aerial vehicle formation with external disturbances and obstacles*

Yang YUAN[1], Yimin DENG[1], Sida LUO[2], Haibin DUAN‡[1,3]

[1]*State Key Laboratory of Virtual Reality Technology and Systems, School of Autonomous Science and Electrical Engineering, Beihang University, Beijing 100083, China*

[2]*School of Mechanical Engineering & Automation, Beihang University, Beijing 100191, China*

[3]*Peng Cheng Laboratory, Shenzhen 518000, China*

E-mail: yyuan@buaa.edu.cn; ymdeng@buaa.edu.cn; s.luo@buaa.edu.cn; hbduan@buaa.edu.cn

Received Dec. 2, 2021; Revision accepted May 4, 2022; Crosschecked May 25, 2022

Abstract: We investigate a distributed game strategy for unmanned aerial vehicle (UAV) formations with external disturbances and obstacles. The strategy is based on a distributed model predictive control (MPC) framework and Levy flight based pigeon inspired optimization (LFPIO). First, we propose a non-singular fast terminal sliding mode observer (NFTSMO) to estimate the influence of a disturbance, and prove that the observer converges in fixed time using a Lyapunov function. Second, we design an obstacle avoidance strategy based on topology reconstruction, by which the UAV can save energy and safely pass obstacles. Third, we establish a distributed MPC framework where each UAV exchanges messages only with its neighbors. Further, the cost function of each UAV is designed, by which the UAV formation problem is transformed into a game problem. Finally, we develop LFPIO and use it to solve the Nash equilibrium. Numerical simulations are conducted, and the efficiency of LFPIO based distributed MPC is verified through comparative simulations.

Key words: Distributed game strategy; Unmanned aerial vehicle (UAV); Distributed model predictive control (MPC); Levy flight based pigeon inspired optimization (LFPIO); Non-singular fast terminal sliding mode observer (NFTSMO); Obstacle avoidance strategy

https://doi.org/10.1631/FITEE.2100559 **CLC number:** TP242.6; V279

1 Introduction

Unmanned aerial vehicles (UAVs) have attracted wide attention with their advantages of low cost, simple operation, and high reliability (Wang B et al., 2020). Compared with manned aerial vehicles (MAVs), UAVs are more suitable for boring, harsh, and dangerous tasks. For these reasons, they are popular in cargo transportation, aerial photography, and other civil and military fields (Labbadi and Cherkaoui, 2019; Zheng and Cai, 2021). In recent years, the UAV swarm has become an important topic in academia and industry (Huo et al., 2021; Li W et al., 2021; Luo et al., 2021). Studies have indicated that the UAV swarm can expand the field of application for UAVs, and has more reliable, more robust, and more durable task execution capabilities (He et al., 2018). Formation control is a key technology for UAVs to realize the collaborative work of a swarm system, and there has already been a wealth of work on the topic. Common formation control methods include the leader–follower strategy, the

‡ Corresponding author

* Project supported by the Science and Technology Innovation 2030 Key Project of "New Generation Artificial Intelligence," China (No. 2018AAA0100803) and the National Natural Science Foundation of China (Nos. T2121003, U1913602, U20B2071, 91948204, and U19B2033)

ORCID: Yang YUAN, https://orcid.org/0000-0002-0715-987X; Yimin DENG, https://orcid.org/0000-0003-1533-3839; Sida LUO, https://orcid.org/0000-0002-5673-6100; Haibin DUAN, https://orcid.org/0000-0002-4926-3202

behavior-based method, and game theory (Dong et al., 2016; Ran et al., 2019; Yang J et al., 2019).

In the leader–follower formation framework, the leader tracks a predetermined trajectory, while the follower interacts with the leader to maintain a relative position. Wang AJ et al. (2018) studied the tracking consistency of the second-order system by designing a fractional observer, and the event-triggered control strategy of periodic sampling was used for the leader–follower formation by using relative position information. However, the structure is not robust enough to guarantee the formation if the leader breaks down or if the communication is unstable. He et al. (2018) proposed a multi-implicit leader formation control algorithm to solve the above problems, adopting a consensus protocol applicable to both the leader and follower. As there was no explicit leader, the damage of a single node had little effect on the stability of the whole swarm. Trinh et al. (2021) studied the bearing-constrained leader–follower formation control problem depending on displacement and bearing vectors, where the relative velocity, bearing rate, and information exchange were not required. Xia et al. (2022) transformed the leader–follower formation problem of heterogeneous systems with time-varying output into a conventional tracking control problem, where a resilient observer was used to eliminate the effect of a cyber-attack.

In the behavior-based method, each UAV has an equal status, and damage to a single UAV node does not affect the normal work of the entire system. Based on the Reynolds rule, Olfati-Saber (2006) proposed three distributed clustering algorithms that can realize a self-organizing formation. Lee G and Chwa (2018) developed a decentralized, behavior-based formation controller considering obstacle avoidance for multiple robots. Qiu and Duan (2020) designed a distributed UAV-formation control framework based on biological behavior, where the UAV control problem was transformed into an object-optimization problem. Tan et al. (2021) proposed the hybrid behavior-based coordination–control method for multiple unmanned surface vehicles, which was effective in a dynamically changing or unknown environment. Liu et al. (2021) studied a behavior-based cooperative target tracking strategy for a dual robotic-dolphin system, and achieved high-level decision-making by the combination of the behavior-based approach and a centralized architecture.

However, the above research considers only the performance of the formation, and the target of a single UAV is ignored. For example, the fuel consumption, preset speed, and attitude requirements of a UAV should be concerned with in some situations. When each UAV has its own objective, UAV formation can be regarded as a game problem in the case of different individual objectives. Gu (2008) proposed a differential game method for the formation control of robots, where the formation control was expressed as a linear quadratic Nash differential game by means of graph theory. Jond and Nabiyev (2019) discussed the coupled Riccati differential equation of the solvability of the differential game method for the formation control problem, and proved the existence of the Nash equilibrium in the discrete formation control problem. Li JQ et al. (2021) studied a distributed game strategy for multi-spacecraft formation control under nonlinear dynamics and perturbation, proposed a worst-case Nash equilibrium strategy, and proved the existence of an open-loop Nash equilibrium. Li YB and Hu (2022) addressed the non-cooperative formation control problem of a multi-agent system by differential game strategy, and studied the Nash equilibria of the finite horizon and infinite horizon games.

Formation control based on a distributed model predictive control (MPC) framework and the optimization algorithm has also been a focus of research in game theory. To solve the formation control problem of UAVs, Zhao et al. (2022) designed a coordinated control scheme based on a distributed MPC. The cost functions of heterogeneous roles such as the leader, coordinator, and follower have been established, and solved by the particle swarm optimization (PSO) algorithm. Yu et al. (2021) defined the tracking task as a distributed MPC problem, and proposed a Nash-combined adaptive differential evolution (ADE) method by combining the ADE algorithm with Nash optimization. Wang YX et al. (2020) proposed a distributed MPC method based on swarm intelligence to solve the local finite time domain optimal control problem using the chaotic gray-wolf optimization (CGWO) method, and adopted an event-triggered strategy to reduce the computation burden.

However, the influence of disturbance on formation was not considered in the above-mentioned research. Keeping the original formation when avoiding obstacles does not guarantee optimal energy for all individuals in large clusters. Based on the above problems, we propose a distributed UAV formation control framework in complex scenarios with external disturbances and obstacles. First, a non-singular fast terminal sliding mode observer (NFTSMO) is proposed to observe the disturbance, and a non-disturbance UAV model is obtained. Second, a distributed UAV obstacle avoidance strategy is proposed with topology reconstruction. Finally, the cost function of each UAV is established in the distributed MPC framework, and a Levy flight based pigeon inspired optimization (LFPIO) algorithm is proposed to solve the distributed cost function to achieve the Nash equilibrium.

The contributions of this study can be summarized as follows: (1) The robustness against disturbance is provided by NFTSMO and fast convergence of the observer is achieved; (2) More energy of a single UAV can be retained during the process of obstacle avoidance by the developed obstacle avoidance strategy; (3) High precision formation with an unknown disturbance can be guaranteed with the proposed NFTSMO and LFPIO based distributed MPC.

2 UAV model and problem formulation

In this study, to design the formation controller of UAVs, we adopt the representative UAV model widely used in much of the literature (Lin, 2014; Wei et al., 2021), in which the dynamic model can be described as

$$\begin{cases} \dot{x}_i = V_i \cos\gamma_i \cos\chi_i, \\ \dot{y}_i = V_i \cos\gamma_i \sin\chi_i, \\ \dot{z}_i = -V_i \sin\gamma_i, \\ \dot{V}_i = \left(T_i - D_i\right)/m_i - g\sin\gamma_i + d_{V_i}, \\ \dot{\gamma}_i = \left(L_i \cos\varphi_i - m_i g\cos\gamma_i\right)/\left(m_i V_i\right) + d_{\gamma_i}, \\ \dot{\chi}_i = \left(L_i \sin\varphi_i\right)/\left(m_i V_i \cos\gamma_i\right) + d_{\chi_i}, \end{cases} \tag{1}$$

where x_i, y_i, and z_i represent the location under the inertial system of coordinates, V_i, χ_i, and γ_i indicate the ground speed, heading angle, and flight path angle, respectively, L_i, D_i, and T_i are lift, drag, and engine thrust, respectively, φ_i is the banking angle, m is the mass of the UAV, g is gravitational acceleration, d_{V_i}, d_{χ_i}, and d_{γ_i} are the external disturbances of V_i, χ_i, and γ_i, respectively, and the control inputs of the model are L_i, T_i, and φ_i.

Defining $p_i=[x_i, y_i, z_i]^{\mathrm{T}}$ and $v_i=\left[\dot{x}_i, \dot{y}_i, \dot{z}_i\right]^{\mathrm{T}}$, and taking the derivative of v_i, we can obtain

$$\dot{v}_i = \Theta_i \tau_i + a_i + d_i, \tag{2}$$

where $\Theta_i = \begin{bmatrix} \cos\chi_i\cos\gamma_i & -\sin\chi_i & -\sin\gamma_i\cos\chi_i \\ \sin\chi_i\cos\gamma_i & \cos\chi_i & -\sin\gamma_i\sin\chi_i \\ \sin\gamma_i & 0 & \cos\gamma_i \end{bmatrix}$, $\tau_i = [\tau_{x_i}, \tau_{y_i}, \tau_{z_i}]^{\mathrm{T}}$, $\tau_{x_i} = (T_i - D_i)/m_i$, $\tau_{y_i} = L_i\sin\varphi_i/m_i$, $\tau_{z_i} = L_i\cos\varphi_i/m_i$, $a_i=[0, 0, -g]^{\mathrm{T}}$, and $d_i=\Theta_i[d_{V_i}, V_i\cos\gamma_i d_{\chi_i}, V_i d_{\gamma_i}]^{\mathrm{T}}$. When we define $u_i=\Theta_i\tau_i+a_i$, the model can be rewritten as

$$\begin{cases} \dot{p}_i = v_i, \\ \dot{v}_i = u_i + d_i. \end{cases} \tag{3}$$

Defining $z_i=[p_i^{\mathrm{T}}, v_i^{\mathrm{T}}]^{\mathrm{T}}$, we can further obtain

$$\dot{z}_i = Az_i + Bu_i + Bd_i, \tag{4}$$

where $A = \begin{bmatrix} 0 & I_3 \\ 0 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 0 \\ I_3 \end{bmatrix}$.

Remark 1 In this paper we focus on formation control and assume that u_i can be ideally controlled.

The schematic of the proposed distributed game strategy is shown in Fig. 1.

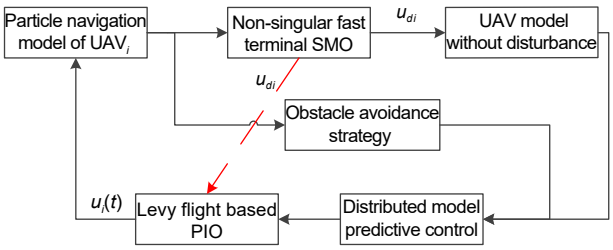

Fig. 1 Schematic of the distributed game strategy

To compensate for the effects of disturbance, we propose an NFTSMO to estimate the disturbance value and then obtain a UAV model without disturbance. By using u_{di} to compensate for the disturbance, the control input u_i can be redefined as $u_i=u_{di}+$

u_{ni}, where u_{ni} is used to achieve the goal of formation. Eq. (4) can be further derived as

$$\begin{aligned}\dot{z}_i &= Az_i + Bu_i + Bd_i = Az_i + B\left(u_{ni} + u_{di}\right) + Bd_i \\ &= Az_i + Bu_{ni} + B\left(u_{di} + d_i\right).\end{aligned} \tag{5}$$

When $u_{di}+d_i$ converges to zero, it can be approximated that $\dot{z}_i = Az_i + Bu_{ni}$.

Graph theory is used to represent the information exchange among UAVs. We define $G=\{V, \varepsilon\}$, where $\mathrm{UAV}_i \in V$ represents a UAV node, and $e_{ij}\in\varepsilon$ indicates the information transmission from UAV_j to UAV_i. The adjacent matrix of G is defined as $H=[h_{ij}]\in\mathbb{R}^{N_{\mathrm{UAV}}\times N_{\mathrm{UAV}}}$, where N_{UAV} is the number of UAVs in the formation, and element h_{ij} is defined as $h_{ij}=\begin{cases}1, & e_{ij}\in\varepsilon,\\ 0, & e_{ij}\notin\varepsilon.\end{cases}$

Remark 2 To guarantee formation consistency, there is at least one directed spanning tree in graph G.

The distributed MPC framework is applied to achieve a formation with the Nash equilibrium strategy. Each UAV is configured with MPC and MPC is applied to determine its own behavior according to the information exchanged, independent of each other. The performance index for each UAV in the framework is designed as

$$\begin{aligned}J_i(t) = &\sum_j h_{ij}\left(\sum_{k=0}^{N-1}\left\| z_i(t+\delta k) - z_j(t+\delta k) - z_{ij}^d \right\|_q\right)\\ &+\sum_{k=0}^{N-1}\left\| u_{ni}(t+\delta k)\right\|_r\\ &+\sum_j h_{ij}\left\| z_i(t+Nk) - z_j(t+Nk) - z_{ij}^d\right\|_m\\ &+\sum_{k=0}^{N}\left\| v_i(t+\delta k) - v_i^d(t+\delta k)\right\|_n,\end{aligned} \tag{6}$$

subject to $z_i\in\mathbb{Z}$, $u_{ni}+u_{di}\in U$, where δ is the time step, z_{ij}^d is the final relative state between UAV_i and UAV_j, which is determined by formation shape z^d, $v_i^d(t)$ is the preset speed to complete the task, q, r, m, and n are the weights of the costs, and N is the number of prediction steps. The performance indices consist of four parts: the desired formation configuration in $0-(N-1)$ steps, the energy consumption in the process, the terminal state penalty function in the N^{th} step, and the penalty function for the predetermined velocity.

The Nash equilibrium strategy of the UAV formation is defined as follows (Lee SM et al., 2015):

$$J_i\left(u_1^*, u_2^*, \cdots, u_{N_{\mathrm{UAV}}}^*\right)\le J_i\left(u_1^*, \cdots, u_{i-1}^*, u_i, u_{i+1}^*, \cdots, u_{N_{\mathrm{UAV}}}^*\right), \tag{7}$$

for i=1, 2, ⋯, N_{UAV}. When the system reaches the Nash equilibrium, no UAV can further optimize its cost by changing its own strategy under the condition that all the other UAVs remain unchanged.

Due to the state variables, and because the control inputs of multiple UAVs are coupled together in the cost function, it is difficult to find a Nash equilibrium strategy to ensure stability by optimizing the control input sequence within a limited prediction range in a distributed MPC framework. Therefore, heuristic optimization algorithms are used to solve the problem, and some studies have shown that evolutionary algorithms are effective. In an LFPIO-based distributed MPC framework, each UAV has its own population to optimize the cost function value, and the local optimization problem is solved with LFPIO according to the Nash equilibrium strategy.

3 Distributed game strategy for UAV formations

3.1 Non-singular fast terminal sliding mode observer

For the particle model of this paper, we consider the system

$$\dot{z} = Az + Bu + d, \tag{8}$$

where $z\in\mathbb{R}^{n\times 1}$ is the state, $A\in\mathbb{R}^{n\times n}$ is the state matrix, and $B\in\mathbb{R}^{n\times m}$ is the input matrix, $u\in\mathbb{R}^{m\times 1}$ is the control input, and $d\in\mathbb{R}^{n\times 1}$ is the external disturbance.

Sliding mode observers with different forms have been developed to solve various problems (Xiong and Saif, 2001; Kalsi et al., 2010; Yang HY et al., 2022). For fast convergence, NFTSMO is developed in this study to achieve high precision formation with unknown disturbances. The non-singular fast terminal sliding surface can be defined as follows (Labbadi and Cherkaoui, 2020; Wang X et al., 2022):

$$s = e + \Pi_1\mathrm{sig}^{\lambda_1}(e) + \Pi_2\mathrm{sig}^{\lambda_2}(\dot{e}), \tag{9}$$

where $e=z-\hat{z}$, $\hat{z}$ is the observed state, λ_1 and λ_2 are positive numbers larger than 1, Π_1=diag(λ_{11}, λ_{12}, ⋯, λ_{1n}), Π_2=diag(λ_{21}, λ_{22}, ⋯, λ_{2n}), and $\lambda_{ij}>0$ (i=1, 2, j=1, 2, ⋯, n).

Motivated by the observer forms designed in previous studies (Kalsi et al., 2010; Czyżniewski and Łangowski, 2022), the disturbance observer is designed as follows:

$$\begin{cases}\dot{\hat{z}}=Az+Bu+s_0+\hat{d},\\ \dot{s}_0=\left(\Pi_2\lambda_2\mathrm{diag}\left(|\dot{e}|^{\lambda_2-1}\right)\right)^{-1}\left[\dot{e}+\Pi_1\lambda_1\mathrm{diag}\left(|e|^{\lambda_1-1}\right)\dot{e}\right],\\ \dot{\hat{d}}=\eta\mathrm{sig}(s)+\rho\mathrm{sig}(s),\end{cases} \tag{10}$$

where $\hat{d}$ is the estimated value of d, η is a positive constant larger than the bound of $\dot{d}$, ρ is a small positive scalar.

We pick the differential of the observed state error as

$$\begin{aligned}\dot{e}&=\dot{z}-\dot{\hat{z}}=(Az+Bu+d)-(Az+Bu+s_0+\hat{d})\\&=d-\hat{d}-s_0.\end{aligned} \tag{11}$$

The derivative of s can be calculated by

$$\begin{aligned}\dot{s}&=\dot{e}+\Pi_1\Gamma_1\mathrm{diag}\left(|e_i|^{\Gamma_1-1}\right)\dot{e}+\Pi_2\Gamma_2\mathrm{diag}\left(|\dot{e}_i|^{\Gamma_2-1}\right)\ddot{e}\\&=\dot{e}+\Pi_1\Gamma_1\mathrm{diag}\left(|e_i|^{\Gamma_1-1}\right)\dot{e}\\&\quad+\Pi_2\Gamma_2\mathrm{diag}\left(|\dot{e}_i|^{\Gamma_2-1}\right)\left(\dot{d}-\dot{\hat{d}}-\dot{s}_0\right)\\&=\Pi_2\Gamma_2\mathrm{diag}\left(|\dot{e}_i|^{\Gamma_2-1}\right)\left(\dot{d}-\dot{\hat{d}}\right).\end{aligned} \tag{12}$$

We define the Lyapunov function as V_o=0.5s^Ts, of which the derivative is

$$\begin{aligned}\dot{V}_o&=s^T\Psi\left(\dot{d}-\eta\mathrm{sig}(s)-\rho\mathrm{sig}(s)\right)\\&\le s^T\Psi\left(l\mathrm{sig}(s)-\eta\mathrm{sig}(s)-\rho\mathrm{sig}(s)\right)\\&\le-\lambda_{\min}(\Psi\xi)\|s\|\le-2\lambda_{\min}(\Psi\xi)V_o,\end{aligned} \tag{13}$$

where $\Psi=\Pi_2\Gamma_2\mathrm{diag}\left(|\dot{e}_i|^{\Gamma_2-1}\right)$ and $\xi=\eta+\rho-l$.

Lemma 1 (Bhat and Bernstein, 2000) Lyapunov function V is defined in the neighborhood Ω ($\Omega\in\mathbb{R}$) of the origin. If $\dot{V}(x)+lV^{\mu}(x)\le0, x\in\Omega$, where $l>0$, $0<\mu<1$, the state x converges to the original neighborhood in a fixed time $T\le\frac{V^{1-\mu}(x_0)}{l(1-\mu)}$, where $V(x_0)$ is the initial value of $V(x)$.

Lemma 1 indicates that NFTSMO is fixed-time convergent. We then divide u into two parts, u_n and u_d, and Eq. (4) is reformed as

$$\dot{z}=Az+Bu_n+Bu_d+Bd. \tag{14}$$

We define $u_d=B^+\hat{d}$, where B^+ is the Moore-Penrose generalized inverse matrix of B, and substitute it into Eq. (14). We can obtain

$$\dot{z}=Az+Bu_n+\hat{d}+Bd. \tag{15}$$

Since $\hat{d}$ converges to Bd in a fixed time, the model is considered as $\dot{z}=Az+Bu_n$ when designing the formation controller.

3.2 Obstacle avoidance strategy

In this study, we propose an obstacle avoidance method based on topology reconstruction. As shown in Fig. 2, a safe circle with radius R_i is set to ensure a safe distance. There are two cases in which the UAV formation encounters obstacles, and the three dashed lines in the figure are in the same direction as the ideal speed of the formation. In the first case, the UAV formation is on one side of the middle, dashed line, and the UAVs avoid the obstacle while keeping formation. In the second case, UAVs are on either side of the middle, dashed line. The UAV formation is split into two parts and flies over the obstacle on either side.

The specific obstacle avoidance strategy can be divided into the following steps:

Step 1: UAV$_i$ judges whether UAV$_j$ is within the obstacle range shown in Fig. 2. If so, go to the next step; otherwise, exit.

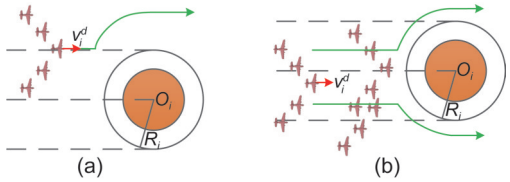

Fig. 2 Two cases in which UAVs encounter obstacles: (a) first case; (b) second case

Step 2: The distance between UAV$_i$ and the obstacle is detected. If the distance is less than d^o, go to the next step; otherwise, exit.

Step 3: A new adjacency relationship with other UAVs based on the position information is established, as shown in Fig. 2b, and the new sub-formation Σ_k is fully connected. The UAVs in the obstacle range in Σ_k are defined as set Σ_{kd}, and UAV$_j$ closest to the obstacle is located.

Step 4: The reference speed $v_j^d(t)$ of UAV$_j$ when avoiding obstacles is calculated. As shown in Fig. 3, the ideal speed of UAV$_j$ is $v_j^d(t)=\left\|v_j(t)\right\|\dfrac{l_1}{\left\|l_1\right\|}$ at point P_1, where l_1 is in the tangent line of UAV$_j$ to the safety circle, and the angle between $v_j(t)$ and l_1 is smaller than the angle between $v_j(t)$ and l_2. As UAV$_j$ moves to point P_2, the tangent line of UAV$_j$ to the safety circle is l_3, and the ideal speed changes as $v_j^d(t)=\left\|v_j(t)\right\|\dfrac{l_3}{\left\|l_3\right\|}$.

Step 5: $v_j^d(t)$ is transmitted to sub-formation Σ_k.

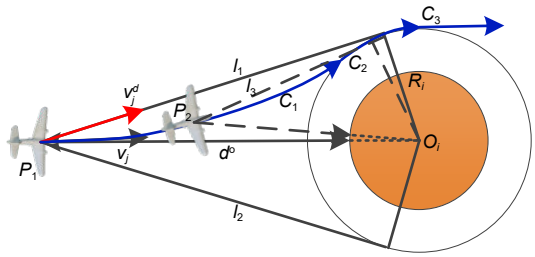

Fig. 3 Ideal speed for obstacle avoidance

The trajectory created by $v_j^d(t)$ can be divided into curves C_1, C_2, and C_3. In the first stage, UAV$_j$ approaches the safety circle along curve C_1 as its velocity $v_j(t)$ is in the right hand of the desired velocity $v_j^d(t)$. In the second stage, UAV$_j$ moves along the safety circle until velocity $v_j(t)$ is in line with the preset velocity. In the final stage, UAV$_j$ successfully avoids the obstacle and moves by the rules of UAV formation.

Due to the strategy of selecting UAV$_j$ in Σ_{kd} closest to the obstacle to generate the obstacle avoidance speed of the sub-formation, the safe distance from the obstacle can be guaranteed. In case 1, the formation remains fully connected during obstacle avoidance. If there is a single individual on one side in case 2, the cost function of the individual retains the control assumption term and velocity term. The preset reference speed is restored when the obstacle is passed.

3.3 Levy flight based pigeon inspired optimization

Pigeon inspired optimization (PIO) is a novel heuristic optimization algorithm proposed by Duan and Qiao (2014), and has been widely used in many fields (Ruan and Duan, 2020). PIO simulates the biological mechanism of the homing navigation of pigeons, and different operators are designed according to navigation tools in different stages. In the first stage, the pigeons use the geomagnetic field and the sun as navigation tools, and a map and compass operator is established. In the second stage, a landmark operator is designed to simulate the effect on pigeons of landmarks near the destination.

We set the swarm size of pigeons to N_p, and the maximum numbers of iterations of the first stage and the second stage are N_c1max and N_c2max, respectively. The initial position of the i^th pigeon is X_0^i, and the velocity is V_0^i. In the first stage, the updating rules of position and speed are as follows:

$$\begin{cases} V_l^i = V_{l-1}^i \mathrm{e}^{-Rl} + \mathrm{rand}\left(X_\mathrm{g} - X_{l-1}^i\right), \\ X_l^i = X_{l-1}^i + V_l^i, \end{cases} \tag{16}$$

where l indicates the generation number, R represents the map and compass factor, X_g is the global best position, and rand is a random number between 0 and 1. When generation l exceeds N_c1max, the updating rules are transformed as follows:

$$\begin{cases} X_l^i = X_{l-1}^i + \mathrm{rand}\,(X_l^\mathrm{c} - X_{l-1}^i), \\ X_l^\mathrm{c} = \sum\limits_{i=1}^{N_\mathrm{p}(l)} X_l^i f\,(X_l^i)/N_\mathrm{p} \sum\limits_{i=1}^{N_\mathrm{p}(l)} f\,(X_l^i), \\ N_\mathrm{p}(l) = \mathrm{ceil}\left(\dfrac{N_\mathrm{p}(l-1)}{2}\right), \end{cases} \tag{17}$$

where $f\,(X_l^i)$ represents the fitness of X_l^i, which is the reciprocal of J_i. In the landmark operator phase, half of the pigeons with poor performance in each generation are eliminated, which can accelerate the convergence of pigeon swarm. However, it is also easy to fall into the local optima (Zhang et al., 2017).

To solve the problem of premature convergence of the population, we propose a Levy flight based PIO, which combines two operators in one stage. The improved updating rules of position and speed are as follows:

$$\begin{cases} V_l^i = V_{l-1}^i \mathrm{e}^{-Rl} + \mathrm{rand}\left(X_\mathrm{g} - X_l^i\right), \\ Y_1^i = X_{l-1}^i + V_l^i, \\ Y_2^i = Y_1^i + r_1 s \oplus \left(X_l^\mathrm{c} - X_{l-1}^i\right), \\ X_l^i = \begin{cases} Y_1^i, f\left(Y_1^i\right) \geq f\left(Y_2^i\right), \\ Y_2^i, f\left(Y_1^i\right) < f\left(Y_2^i\right), \end{cases} \\ X_l^\mathrm{c} = \sum_{i=1}^{N_\mathrm{p}} X_l^i f\left(X_l^i\right) / N_\mathrm{p} \sum_{i=1}^{N_\mathrm{p}} f\left(X_l^i\right), \end{cases} \tag{18}$$

where r_1 is the scaling factor, s is the Levy flight operator, and $\oplus$ is the dot product operator. The Levy flight model is an optimal random walk model with a heavy tail probability distribution (Feng et al., 2021), which can be described as follows:

$$\begin{cases} s = \mu / |v|^{\frac{1}{\delta}}, \\ \mu \sim N\left(0, \sigma_\mu^2\right), v \sim N\left(0, \sigma_v^2\right), \\ \sigma_\mu = \left\{ \dfrac{\Gamma(1+\delta)\sin(\pi\delta/2)}{\Gamma[(1+\delta)/2]\,\delta \cdot 2^{(\delta-1)/2}} \right\}^{\frac{1}{\delta}}, \end{cases} \tag{19}$$

where $\delta \in (0, 2]$ is a constant, σ_v=1, and $N(0, \sigma^2)$ represents the normal distribution.

3.4 LFPIO based Nash equilibrium strategy for distributed MPC

There is a separate cost function for each UAV, and a separate pigeon flock S_i is designed for each cost function for optimization. The control input sequence at time t obtained by solving the distributed MPC can be expressed as $U_{ni}(t)=[u_{ni}(0|t), \cdots, u_{ni}(N_\mathrm{UAV}-1|t)]$, and $u_{ni}(0|t)$ is used as the control input for the UAV model. We define $S_i X_l^j = [u_{ni}^j(0|t), \cdots, u_{ni}^j(N_\mathrm{UAV}-1|t)]$, $j \in \{1, 2, \cdots, N_\mathrm{p}\}$, which represents the position of the j^{th} pigeon at generation l in swarm S_i corresponding to UAV$_i$, and the velocity of the j^{th} pigeon is $S_i V_l^j$. The solution process is shown in Fig. 4.

Step 1: The control input sequence $U_{ni}(t-\delta)$ at the previous moment is inherited at time t, and it is used as the initial input sequence of the current moment after the following processing:

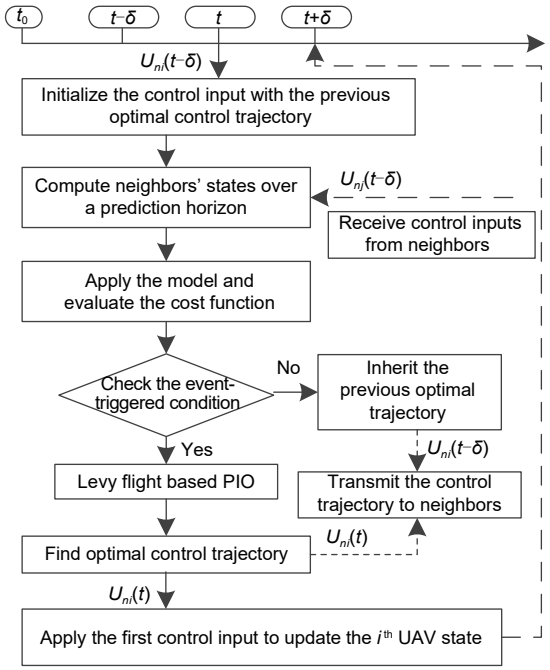

Fig. 4 LFPIO based Nash equilibrium strategy for distributed MPC

$$\begin{aligned} U_{ni}(t) = \big[& u_{ni}\left(1|t-\delta\right), \cdots, u_{ni}\left(N_\mathrm{UAV}-1|t-\delta\right), \\ & u_{ni}\left(N_\mathrm{UAV}-1|t-\delta\right) \big]. \end{aligned} \tag{20}$$

It can be seen that $U_{ni}(t)$ inherits the predicted value of $U_{ni}(t-\delta)$ from step 1 to step $N-1$, and the predicted value at step $N-1$ is also assigned to $u_{ni}(N_\mathrm{UAV}-1|t)$.

Step 2: The received neighbor information is processed according to step 1, and the cost function of UAV$_i$ is calculated. We determine whether the cost $J_i(t)$ is greater than the triggered threshold T_J. If it is, the current input does not meet the requirement, and LFPIO is triggered to optimize the input sequence. If not, $U_{ni}(t)$ will be output.

Step 3: Using LFPIO to optimize the cost function, $U_{ni}(t)$ from step 1 is taken as the initial global optimal position. The updated global optimal position is output as the control input sequence $U_{ni}(t)$ after optimization.

Step 4: The output sequence $U_{ni}(t)$ is sent to the neighbor and executed as the first control input in the sequence.

4 Simulation results

In this section, the numerical simulation conducted for the two cases is shown. The simulation step δ was 0.2 s, and the prediction horizon of MPC was $[t, t+4\delta]$. A swarm of five UAVs was used for the simulation, of which the initial speeds were all set as 20 m/s, and the position of each UAV was generated randomly. In both cases, the UAVs were subject to external interference, and the disturbance of UAV_i was set as

$$d_i = \begin{bmatrix} (-1)^i 0.2\sin(\pi t/20) \\ (-1)^i 0.2\cos(\pi t/20) \\ (-1)^i 0.2\sin(\pi t/20) \end{bmatrix} + 0.1\,\mathrm{rand}. \tag{21}$$

The initial adjacent matrix and formation shape z^d were set as

$$H = \begin{bmatrix} 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 \end{bmatrix},$$
$$z^d = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ -10 & 20 & 0 & 0 & 0 & 0 \\ -10 & -20 & 0 & 0 & 0 & 0 \\ -20 & 40 & 0 & 0 & 0 & 0 \\ -20 & -40 & 0 & 0 & 0 & 0 \end{bmatrix}^{\mathrm{T}}, \tag{22}$$

and z_{ij}^d is the result of the j^{th} column of z^d minus the i^{th} column.

Four different methods were used to generate control inputs, PIO (PIO without using NFTSMO), LFPIO (LFPIO without using NFTSMO), PIO with NFTSMO, and LFPIO with NFTSMO. The parameters of PIO and LFPIO are given in Table 1.

Table 1 Parameters of PIO and LFPIO

Algorithm	Description	Value
PIO	Maximum number of iterations of operator 1 (N_{c1max})	25
	Maximum number of iterations of operator 2 (N_{c2max})	5
	Map and compass operator (R)	0.1
LFPIO	Maximum number of iterations of LFPIO (N_{cmax})	30
	Map and compass operator (R)	0.1
	Scaling factor (r_1)	0.01

To reduce the computational burden, the total number of iterations of PIO and LFPIO was set as 30. R was set to 0.1 to obtain a good search capability. Setting r_1 to 0.01 increased its development capability after exploration, and the ability was poor when r_1 was large.

The desired speed was set as

$$V_i^d = \begin{cases} [20, 0, 0]^{\mathrm{T}}, & 0 \le t \le 50, \\ [0, 20, 0]^{\mathrm{T}}, & 50 < t \le 100, \\ [-20, 0, 0]^{\mathrm{T}}, & 100 < t \le 150, \\ [0, -20, 0]^{\mathrm{T}}, & 150 < t \le 200. \end{cases} \tag{23}$$

There were two obstacles in the flight path of the UAVs, of which the coordinates were $(300, 0)^{\mathrm{T}}$ and $(1050, 400)^{\mathrm{T}}$, and the radius of the obstacles was 30 m.

The simulation results of the trajectories are shown in Fig. 5. First, the UAV formation can successfully complete the task and navigate through all obstacles with the obstacle avoidance strategy proposed in this study. When crossing the first obstacle (Fig. 2b), the UAVs were automatically divided into two subformations. When facing the second obstacle (Fig. 2a), the UAVs chose to keep the formation, flying from one side. The effectiveness of the obstacle avoidance strategy proposed is clearly shown. Second, the ideal

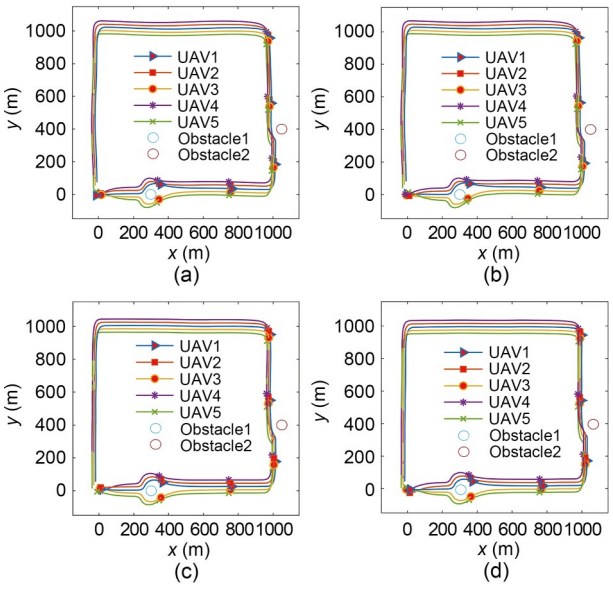

Fig. 5 Simulation results of trajectories: (a) PIO; (b) LFPIO; (c) PIO with NFTSMO; (d) LFPIO with NFTSMO

trajectory generated in each time period with the desired velocity was a straight line in the case of no obstacles. However, we can see that the velocity directions of the UAVs in Figs. 5a and 5b deviated significantly from the predetermined one with the influence of a disturbance. It is apparent that the methods with NFTSMO had better performance than the others. The simulation results in Figs. 5c and 5d were straighter than those in Figs. 5a and 5b. The observation results of the disturbance are shown in Fig. 6, in which the observed value and the actual value were essentially the same in three channels.

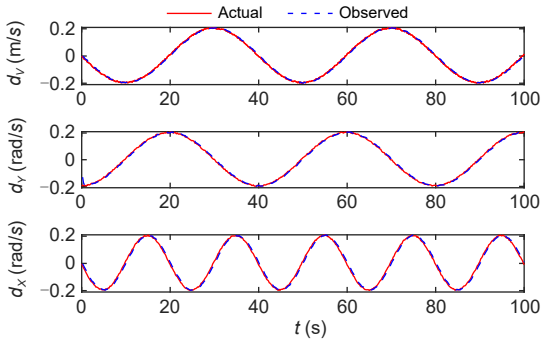

Fig. 6 Observed values of disturbance

The velocities of UAV1 with different methods are shown in Fig. 7. The velocity trends obtained by the four methods were essentially consistent. At the beginning of the curves in Figs. 7a and 7b, the velocity obtained by PIO fluctuated the most, followed by LFPIO and PIO with NFTSMO, and finally LFPIO with NFTSMO. When encountering an obstacle, the speed of UAV1 showed a strong variation. As depicted in the local figure, the speed obtained by LFPIO with NFTSMO was closer to the reference most of the time. The PIO and LFPIO methods had greater velocity deviations than PIO with NFTSMO and LFPIO with NFTSMO, which demonstrates the effectiveness of NFTSMO.

We averaged the cost function for each UAV to obtain J_{mean}, and took log base 10 of J_{mean} to obtain Fig. 8. At the initial moment, the values of J_{mean} of all the four methods were large due to the large difference between the relative position and the ideal relative value. After a period of time, the mean cost value stabilized within a certain range. The UAVs detected obstacles and carried out topology reconstruction at 10 s, and the cost function increased suddenly due to

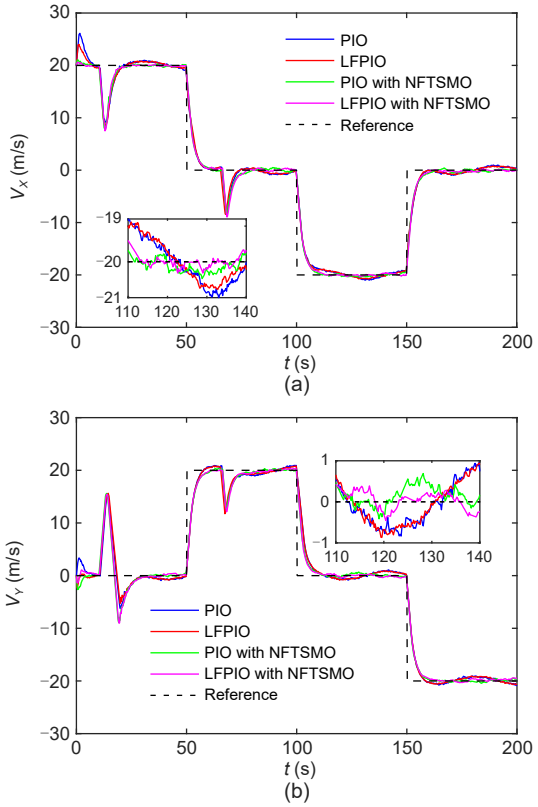

Fig. 7 Velocity of UAV1 with different methods: (a) velocity along the *X* axis; (b) velocity along the *Y* axis

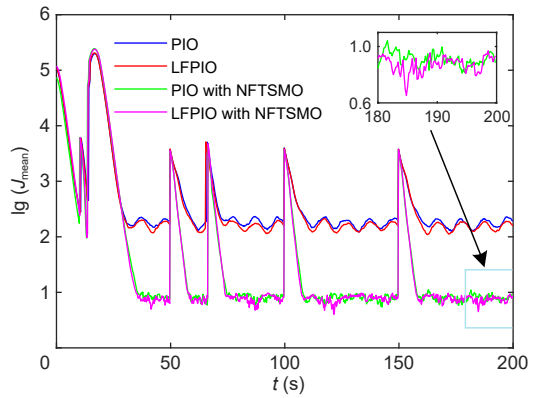

Fig. 8 Average cost with different methods

the change of the UAVs' adjacent matrix. The UAVs completed the obstacle avoidance, and the formation topology was restored to the initial state at 13.6 s. At the time, the UAVs were located on both sides of the obstacle and the distance between the two subgroups was the largest, thus the cost function mutated again. The cost function increased suddenly as the UAVs' preset speed changed at 50, 100, and 150 s. The cost function changed suddenly at 66.2 s because the

UAVs encountered obstacles again. We can also see that LFPIO performed better than PIO most of time with or without NFTSMO, indicating that LFPIO has stronger search ability than PIO and can acquire a better solution.

Fig. 9 shows the triggered times of UAV5 with different methods. The triggered threshold was set as 10, indicating that it will not be triggered in the case of $J_i(t)\leq 10$. The effects of the disturbance cannot be ignored, for which PIO and LFPIO were still triggered at every moment. Corresponding to Fig. 5, the control input updated less without encountering obstacles and changing direction, and the triggered times were reduced (Fig. 9).

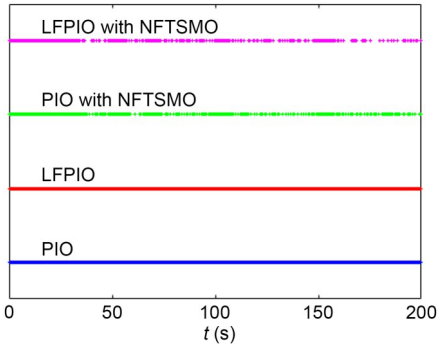

Fig. 9 Triggered times of UAV5 with different methods

The details of triggered times of UAVs with different methods are shown in Table 2. The UAVs need to calculate the control input at every moment in the absence of NFTSMO. The triggered time of LFPIO with NFTSMO was only 455 for UAV1, and the average time was only 498.6, which means that the computing burden was halved. Compared with PIO with NFTSMO, the triggered times of LFPIO with NFTSMO were reduced by 12.8%.

Table 2 Triggered times of UAVs with different methods

Algorithm	Triggered times					
	UAV1	UAV2	UAV3	UAV4	UAV5	Mean
PIO	1000	1000	1000	1000	1000	1000
LFPIO	1000	1000	1000	1000	1000	1000
PIO with NFTSMO	513	632	578	598	539	572
LFPIO with NFTSMO	455	560	501	518	459	498.6

Through the above analysis, we can see that the obstacle avoidance strategy and disturbance observer were very effective, and the optimization ability of LFPIO was greatly improved compared with PIO.

5 Conclusions

In this study, considering the problem of distributed UAV formation, we designed a non-singular fast terminal sliding mode observer to observe the influence of disturbance, and feedforward compensation was conducted to obtain a non-disturbance model. When the UAVs encountered obstacles, topology reconstruction was used to ensure that each UAV can take a small output to avoid obstacles. Based on the above work, a cost function was established in the distributed model predictive control framework, and a Nash equilibrium strategy was adopted. Then, the original pigeon inspired optimization was improved, and a Nash equilibrium can be obtained by Levy flight based pigeon inspired optimization. The simulation results showed that the distributed game strategy is effective.

Contributors

Haibin DUAN and Yang YUAN designed the research. Yang YUAN and Yimin DENG processed the data. Yang YUAN drafted the paper. Sida LUO helped organize the paper. Haibin DUAN and Yang YUAN revised and finalized the paper.

Compliance with ethics guidelines

Yang YUAN, Yimin DENG, Sida LUO, and Haibin DUAN declare that they have no conflict of interest.

References

Bhat SP, Bernstein DS, 2000. Finite-time stability of continuous autonomous systems. *SIAM J Contr Optim*, 38(3): 751-766. https://doi.org/10.1137/S0363012997321358

Czyżniewski M, Łangowski R, 2022. A robust sliding mode observer for non-linear uncertain biochemical systems. *ISA Trans*, 123:25-45. https://doi.org/10.1016/j.isatra.2021.05.040

Dong LF, Chen YZ, Qu XJ, 2016. Formation control strategy for nonholonomic intelligent vehicles based on virtual structure and consensus approach. *Proc Eng*, 137:415-424. https://doi.org/10.1016/j.proeng.2016.01.276

Duan HB, Qiao PX, 2014. Pigeon-inspired optimization: a new swarm intelligence optimizer for air robot path planning. *Int J Intell Comput Cybern*, 7(1):24-37. https://doi.org/10.1108/IJICC-02-2014-0005

Feng X, Muramatsu H, Katsura S, 2021. Differential evolutionary algorithm with local search for the adaptive periodic-

disturbance observer adjustment. *IEEE Trans Ind Electron*, 68(12):12504-12512. https://doi.org/10.1109/TIE.2020.3040664

Gu DB, 2008. A differential game approach to formation control. *IEEE Trans Contr Syst Technol*, 16(1):85-93. https://doi.org/10.1109/TCST.2007.899732

He LL, Bai P, Liang XL, et al., 2018. Feedback formation control of UAV swarm with multiple implicit leaders. *Aerosp Sci Technol*, 72:327-334. https://doi.org/10.1016/j.ast.2017.11.020

Huo MZ, Duan HB, Fan YM, 2021. Pigeon-inspired circular formation control for multi-UAV system with limited target information. *Guid Navig Contr*, 1(1):2150004. https://doi.org/10.1142/S2737480721500047

Jond HB, Nabiyev V, 2019. On the finite horizon Nash equilibrium solution in the differential game approach to formation control. *J Syst Eng Electron*, 30(6):1233-1242. https://doi.org/10.21629/JSEE.2019.06.17

Kalsi K, Lian JM, Hui SF, et al., 2010. Sliding-mode observers for systems with unknown inputs: a high-gain approach. *Automatica*, 46(2):347-353. https://doi.org/10.1016/j.automatica.2009.10.040

Labbadi M, Cherkaoui M, 2019. Robust adaptive backstepping fast terminal sliding mode controller for uncertain quadrotor UAV. *Aerosp Sci Technol*, 93:105306. https://doi.org/10.1016/j.ast.2019.105306

Labbadi M, Cherkaoui M, 2020. Robust adaptive nonsingular fast terminal sliding-mode tracking control for an uncertain quadrotor UAV subjected to disturbances. *ISA Trans*, 99:290-304. https://doi.org/10.1016/j.isatra.2019.10.012

Lee G, Chwa D, 2018. Decentralized behavior-based formation control of multiple robots considering obstacle avoidance. *Intel Serv Robot*, 11(1):127-138. https://doi.org/10.1007/s11370-017-0240-y

Lee SM, Kim H, Myung H, et al., 2015. Cooperative coevolutionary algorithm-based model predictive control guaranteeing stability of multirobot formation. *IEEE Trans Contr Syst Technol*, 23(1):37-51. https://doi.org/10.1109/TCST.2014.2312324

Li JQ, Chen S, Li CY, et al., 2021. Distributed game strategy for formation flying of multiple spacecraft with disturbance rejection. *IEEE Trans Aerosp Electron Syst*, 57(1):119-128. https://doi.org/10.1109/TAES.2020.3010593

Li W, Yang BW, Song GH, et al., 2021. Dynamic value iteration networks for the planning of rapidly changing UAV swarms. *Front Inform Technol Electron Eng*, 22(5):687-696. https://doi.org/10.1631/FITEE.1900712

Li YB, Hu XM, 2022. A differential game approach to intrinsic formation control. *Automatica*, 136:110077. https://doi.org/10.1016/j.automatica.2021.110077

Lin W, 2014. Distributed UAV formation control using differential game approach. *Aerosp Sci Technol*, 35:54-62. https://doi.org/10.1016/j.ast.2014.02.004

Liu JC, Wu ZX, Yu JZ, et al., 2021. Cooperative target tracking in aquatic environment using dual robotic dolphins. *IEEE Trans Syst Man Cybern Syst*, 51(8):4782-4792. https://doi.org/10.1109/TSMC.2019.2944753

Luo YH, Bai A, Zhang HG, 2021. Distributed formation control of UAVs for circumnavigating a moving target in three-dimensional space. *Guid Navig Contr*, 1(3):2150014. https://doi.org/10.1142/S273748072150014X

Olfati-Saber R, 2006. Flocking for multi-agent dynamic systems: algorithms and theory. *IEEE Trans Autom Contr*, 51(3):401-420. https://doi.org/10.1109/TAC.2005.864190

Qiu HX, Duan HB, 2020. A multi-objective pigeon-inspired optimization approach to UAV distributed flocking among obstacles. *Inform Sci*, 509:515-529. https://doi.org/10.1016/j.ins.2018.06.061

Ran MP, Xie LH, Li JC, 2019. Time-varying formation tracking for uncertain second-order nonlinear multi-agent systems. *Front Inform Technol Electron Eng*, 20(1):76-87. https://doi.org/10.1631/FITEE.1800557

Ruan WY, Duan HB, 2020. Multi-UAV obstacle avoidance control via multi-objective social learning pigeon-inspired optimization. *Front Inform Technol Electron Eng*, 21(5):740-748. https://doi.org/10.1631/FITEE.2000066

Tan GG, Zhuang JY, Zou J, et al., 2021. Coordination control for multiple unmanned surface vehicles using hybrid behavior-based method. *Ocean Eng*, 232:109147. https://doi.org/10.1016/j.oceaneng.2021.109147

Trinh MH, van Tran Q, van Vu D, et al., 2021. Robust tracking control of bearing-constrained leader-follower formation. *Automatica*, 131:109733. https://doi.org/10.1016/j.automatica.2021.109733

Wang AJ, Liao XF, Dong T, 2018. Fractional-order follower observer design for tracking consensus in second-order leader multi-agent systems: periodic sampled-based event-triggered control. *J Franklin Inst*, 355(11):4618- 4628. https://doi.org/10.1016/j.jfranklin.2018.01.036

Wang B, Shen YY, Zhang YM, 2020. Active fault-tolerant control for a quadrotor helicopter against actuator faults and model uncertainties. *Aerosp Sci Technol*, 99:105745. https://doi.org/10.1016/j.ast.2020.105745

Wang X, Xu B, Cheng YX, et al., 2022. Robust adaptive learning control of space robot for target capturing using neural network. *IEEE Trans Neur Netw Learn Syst*, early access. https://doi.org/10.1109/TNNLS.2022.3144569

Wang YX, Zhang T, Cai ZH, et al., 2020. Multi-UAV coordination control by chaotic grey wolf optimization based distributed MPC with event-triggered strategy. *Chin J Aeronaut*, 33(11):2877-2897. https://doi.org/10.1016/j.cja.2020.04.028

Wei LL, Chen M, Li T, 2021. Disturbance-observer-based formation-containment control for UAVs via distributed adaptive event-triggered mechanisms. *J Franklin Inst*, 358(10):5305-5333. https://doi.org/10.1016/j.jfranklin.2021.04.050

Xia LN, Li Q, Song RZ, et al., 2022. Leader-follower time-varying output formation control of heterogeneous systems under cyber attack with active leader. *Inform Sci*, 585:24-40. https://doi.org/10.1016/j.ins.2021.11.026

Xiong Y, Saif M, 2001. Sliding mode observer for nonlinear uncertain systems. *IEEE Trans Autom Contr*, 46(12):2012-2017. https://doi.org/10.1109/9.975511

Yang HY, Yin S, Han HG, et al., 2022. Sparse actuator and sensor attacks reconstruction for linear cyber-physical systems with sliding mode observer. *IEEE Trans Ind Inform*, 18(6):3873-3884. https://doi.org/10.1109/TII.2021.3111221

Yang J, Wang XM, Baldi S, et al., 2019. A software-in-the-loop implementation of adaptive formation control for fixed-wing UAVs. *IEEE/CAA J Autom Sin*, 6(5):1230-1239. https://doi.org/10.1109/JAS.2019.1911702

Yu Y, Wang HL, Liu SM, et al., 2021. Distributed multi-agent target tracking: a Nash-combined adaptive differential evolution method for UAV systems. *IEEE Trans Veh Technol*, 70(8):8122-8133. https://doi.org/10.1109/TVT.2021.3091575

Zhang DF, Duan HB, Yang YJ, 2017. Active disturbance rejection control for small unmanned helicopters via Levy flight-based pigeon-inspired optimization. *Aircraft Eng Aerosp Technol*, 89(6):946-952. https://doi.org/10.1108/AEAT-05-2016-0065

Zhao J, Sun JM, Cai ZH, et al., 2022. Distributed coordinated control scheme of UAV swarm based on heterogeneous roles. *Chin J Aeronaut*, 35(1):81-97. https://doi.org/10.1016/j.cja.2021.01.014

Zheng Z, Cai SC, 2021. A collaborative target tracking algorithm for multiple UAVs with inferior tracking capabilities. *Front Inform Technol Electron Eng*, 22(10):1334-1350. https://doi.org/10.1631/FITEE.2000362