

【引用格式】董睿冲, 徐涛, 王晓丽. 基于数据驱动的无人机编队最优控制算法设计[J]. 数字海洋与水下攻防, 2023, 6(6): 670-680.

# 基于数据驱动的无人机编队最优控制算法设计

董睿冲, 徐涛, 王晓丽\*

(哈尔滨工业大学(威海)信息科学与工程学院, 山东 威海 264209)

**摘要** 无人机编队具有广泛的应用价值和发展前景, 然而, 无人机动力学模型的非线性和耦合性, 以及实际应用中的不确定性限制了其应用。为了解决模型不精确问题, 提出了一种数据驱动的自适应动态编程(ADP)最优编队控制器设计方法。该方法结合了分布式一致性和线性二次方相关理论。首先, 通过建立虚拟领航-跟随者模型, 可以将四旋翼无人机的复杂非线性动力学模型简化为4个线性子系统, 并解耦它们的相互影响; 接下来, 针对每个子系统设计编队控制器; 随后应用李雅普诺夫理论对算法进行稳定性分析; 最后通过仿真, 验证了所提算法的稳定性和有效性。

**关键词** 无人机编队; 数据驱动; 领航-跟随者; 自适应动态规划

**中图分类号** TP273+1 **文献标识码** A **文章编号** 2096-5753(2023)06-0670-11

**DOI** 10.19838/j.issn.2096-5753.2023.06.004

## Design of Optimal Data-driven Control Algorithm for UAV Formation

DONG Ruichong, XU Tao, WANG Xiaoli\*

(School of Information Science and Engineering, Harbin Institute of Technology, Weihai 264209, China)

**Abstract** UAV formation has a wide range of application value and development prospects. However, the nonlinear and coupling nature of UAV dynamic model and the uncertainties in practical applications limit its application. In order to solve the problem of model inaccuracy, an optimal formation controller design method based on data-driven adaptive dynamic programming (ADP) is proposed. It combines the theory of distributed consistency and linear quadratic correlation. First, a virtual leader-follower model is established, the UAV's nonlinear dynamics model is simplified into four linear subsystems, and their mutual influences are decoupled. Next, the formation controller is designed for each subsystem. Then, the stability of the algorithm is analyzed using Lyapunov stability theory. Finally, the stability and effectiveness of the proposed algorithm is verified.

**Key words** UAV formation; data-driven; leader-follower model; adaptive dynamic programming

## 0 引言

最近几年来, 无人机技术的迅猛发展使得固定

翼无人机和旋翼无人机在各个领域得到了广泛的应用<sup>[1]</sup>。无人机的海陆空协同作战在海洋作战中具有多个优势。首先, 无人机能够在复杂的海洋

收稿日期: 2023-09-29

作者简介: 董睿冲(2003-), 男, 主要从事多智能体协同控制研究。

\*通信作者: 王晓丽(1981-)女, 博士, 教授, 主要从事多智能体协同控制研究。

基金项目: 山东省自然科学基金“复杂多航空器系统分布式控制”(ZR2020MF072); 山东省高校青年创新科技计划“复杂多智能体系统协同控制”(2019KJN004)。

环境中迅速响应和部署;其次,无人机能够通过高空俯瞰和远距离侦察,提供全面的情报支持,提升作战主体的感知能力和对战场情况的认知;此外,无人机在海洋作战中能够执行风险较高的任务,减少人员伤亡风险,提高作战效能和安全性。无人机编队的协同作战将在海洋作战中发挥重要作用,为作战指挥和决策提供更多的选择和灵活性。多无人机编队控制问题一直是无人机领域中备受关注的研究方向,国内外众多学者投入了大量的研究工作。南京信息工程大学自动化学院庄亚楠团队在有向拓扑结构下,针对无人机之间存在通信噪声干扰,研究了基于事件触发的多无人机群快速编队控制<sup>[2]</sup>。南京邮电大学的 LEI JI 教授团队研究了通信带宽有限情况下多无人机系统的编队控制问题,利用位置和速度的组合误差,协同设计了具有事件触发机制的一致性控制协议,以减少网络上的信息传输量<sup>[3]</sup>。在国外,俄罗斯科学院 Pereslavl-Zalessky 程序系统研究所的学者论述并考虑了在不稳定环境下运行的一组无人机编队控制,提出了一种基于自适应 Kohonen 神经网络的方法,应用了智能几何控制的原理,为在复杂环境中实现无人机群的运动提供了理论支撑<sup>[4]</sup>。沙特阿卜杜勒阿齐兹国王大学航空航天工程系的学者提出了一种新的二阶非线性多智能体系统(MAS)的分布式共识算法用于解决多架无人机协同控制的问题<sup>[5]</sup>。

领航-跟随法是无人机编队控制领域的一个研究重点,具有广泛的应用前景。该研究方向旨在实现一个或多个无人机作为领航者,引导其他无人机进行协同飞行。国内外许多学者在这一领域开展了深入的研究,并提出了各类创新方法和技术。在文献[6]中,将旋翼无人机设定为跟随者,地面小车设定为领航者,并采用线性二次调节器(LQR)方法和滑模控制设计了控制器,成功实现了编队控制。另一方面,在文献[7]中,基于虚拟领航-跟随者模型,提出了一种创新的反步控制方法,能够快速实现编队状态的稳定。在文献[8]中,研究者在领航-跟随者结构下,通过设计适当的滑模面和滑

模控制律,能够实现对群体系统的集中控制,使得每个个体能够按照预定的规则和策略行动。滑模控制方法的稳定性和鲁棒性使得群体控制可以在各种环境和条件下有效工作,具有一定的鲁棒性和适应性<sup>[9]</sup>。

虽然目前该领域成果丰硕,但现有研究仍存在一些问題。其一为对无人机进行建模时缺乏足够的精确性。目前大部分研究在建立模型或线性化运动学和动力学模型的过程中。由于简化程度较高,无法完全考虑许多不确定因素,实际控制过程中模型存在不确定性,这使得控制精度的进一步提高受到限制。

数据驱动控制(Data-driven Control)是一种在控制器设计中不依赖于受控系统的数学模型信息,而是仅使用系统的离线/在线数据和经过数据处理得到的信息的方法。通过严格的数学分析,其稳定性和收敛性可以保证其鲁棒性<sup>[10]</sup>。为了解决无人机编队模型建立不准确的问题<sup>[11-12]</sup>,使用数据驱动的方法来控制无人机和编队被证明是有效的,具有实际重要意义<sup>[13-14]</sup>。

本文提出了针对控制器设计过程中需要求解一个复杂的代数黎卡提方程的问题,提出一种基于数据驱动的算法,近似得到黎卡提方程(ARE)唯一解。同时,本文针对无人机编队中由于内部参数摄动导致的模型不精确问题并提出了数据驱动的方法。在适用于明确已知的线性系统模型的同时,只需要一段时间内的输入和状态信息,即可求解出最优的反馈控制策略。

## 1 无人机系统模型

本文研究多无人机编队系统在实际应用中的问题,例如无人机集群对抗和多无人机协同搜索。研究对象为十字形四旋翼无人机。基于虚拟领航-跟随者模型,考虑到存在系统参数摄动的情况下,论文探讨了由多个四旋翼无人机组成的编队最优控制问题。四旋翼无人机是强耦合和欠驱动系统,它有6个自由度。

常见的四旋翼无人机非线性动力学模型被表示为



平衡时固定的控制输入:

$$\bar{U} = [mg \ 0 \ 0 \ 0]^T \in \mathbb{R}^4 \quad (9)$$

这个输入量所代表的为无人机的 4 个转子产生的用于抵消无人机自身重力的上升力, 它使无人机平稳地悬停在规定的平衡点处。从而可以得知该无人机的线性动力学模型的状态空间表达式为

$$\dot{X} = AX + BU \quad (10)$$

式中:

$$A = \left. \frac{\partial f(X, U)}{\partial X} \right|_{\substack{X=\bar{X} \\ U=\bar{U}}} = \begin{bmatrix} 0_{6 \times 6} & I_{6 \times 6} \\ 0 & 0 & 0 & 0 & g & 0 \\ 0 & 0 & 0 & -g & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (11)$$

$$B = \left. \frac{\partial f(X, U)}{\partial U} \right|_{\substack{X=\bar{X} \\ U=\bar{U}}} = \begin{bmatrix} 0_{8 \times 4} \\ \frac{1}{m} & 0 & 0 & 0 \\ 0 & \frac{l}{J_x} & 0 & 0 \\ 0 & 0 & \frac{l}{J_y} & 0 \\ 0 & 0 & 0 & \frac{1}{J_z} \end{bmatrix} \quad (12)$$

由式 (10) 可得, 状态  $x$  和  $y$  仅与姿态角  $\beta, \alpha$  相关, 与控制输入  $U_1$  无关。以此简化后的系统如下:

$$\begin{cases} \dot{x} = x^{(1)} \\ \dot{x}^{(1)} = x^{(2)} \\ \dot{x}^{(2)} = x^{(3)} \\ \dot{x}^{(3)} = U_3' \end{cases} \quad (13)$$

$$\begin{cases} \dot{y} = y^{(1)} \\ \dot{y}^{(1)} = y^{(2)} \\ \dot{y}^{(2)} = y^{(3)} \\ \dot{y}^{(3)} = U_2' \end{cases} \quad (14)$$

$$\begin{cases} \dot{z} = z^{(1)} \\ \dot{z}^{(1)} = U_1' \end{cases} \quad (15)$$

$$\begin{cases} \dot{\gamma} = \gamma^{(1)} \\ \dot{\gamma}^{(1)} = U_4' \end{cases} \quad (16)$$

式中:  $U_3' = glU_3/J_y$ ;  $U_2' = -glU_2/J_x$ ;  $U_1' = U_1/m - g$ ;

$$U_4' = U_4/J_z。$$

通过解耦简化的 4 个线性子系统, 我们可以观察到根据卡尔曼提出的系统能控能观性判据, 这些线性子系统都具备能够被控制和被观察的性质。

## 2.2 编队线性模型

根据 3.1 的内容, 对于无人机的编队控制器, 需要对 4 个子系统单独设计。下面以位置子系统  $x$  为例, 第  $i$  架无人机的数学模型为

$$\dot{x}_i = Ax_i + Bu_i \quad (i=1, 2, 3, 4) \quad (17)$$

式中:

$$x_i \in \mathbb{R}^{4 \times 1}, u_i = U_3', A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}, B = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

这架无人机与理想状态间的误差表示为

$$\zeta_i = (x_i) - \rho_0 - \rho_{i0} \quad (18)$$

式中:  $\rho_0$  表示无人机的期望飞行状态;  $\rho_{i0}$  表示第  $i$  架无人机与期望飞行状态之间的差值, 可将其假设成  $[x_{i0} \ v_{i0} \ 0 \ 0]^T$ 。

对误差变量进行求导可得:

$$\begin{aligned} \dot{\zeta}_i &= (\dot{x}_i) - \dot{\rho}_0 - \dot{\rho}_{i0} \\ &= Ax_i + Bu_i - A\rho_0 - A\rho_{i0} \\ &= A\zeta_i + Bu_i \end{aligned} \quad (19)$$

以无人机编队为整体来说:

$$\begin{aligned} \dot{\zeta} &= (I_N \otimes A)\zeta + (I_N \otimes B)u \\ \zeta &= [\zeta_1^T \ \zeta_2^T \ \dots \ \zeta_N^T]^T \in \mathbb{R}^{nN} \\ u &= [u_1^T \ u_2^T \ \dots \ u_N^T]^T \in \mathbb{R}^{mN} \end{aligned} \quad (20)$$

式中:  $N$  表示编队中受控制无人机数量;  $n$  表示控制对象状态变量的维数; 输入变量的维度为  $m$ , 状态子系统位置  $x$  中,  $m=1, n=4$ 。

在实际应用中, 无人机的负荷与理想情况存在差异, 有些时候无人机没法查清负荷的情况。所以模型中的质量和转动惯量不是已知的。基于这些情况, 我们可以改写编队的误差状态空间如下:

$$\dot{\zeta} = (I_N \otimes \hat{A})\zeta + (I_N \otimes \hat{B})u \quad (21)$$

式中:  $\hat{A}, \hat{B}$  为状态矩阵和输入矩阵, 但是有如下未

知参数:

$$\hat{A} = A + \Delta A, \hat{B} = B + \Delta B \quad (22)$$

$\Delta A, \Delta B$  为摄动的动力学参数, 定义范数有界的不确定参数:

$$\Delta A = G_1 M_1 E_1, \Delta B = G_2 M_2 E_2 \quad (23)$$

式中:  $G_i, E_i (i=1,2)$  是由常数组成的矩阵,  $M_i (i=1,2)$  是一个有界未知的矩阵且满足:

$$M_i = \{M | M^T M \leq I\} \quad (24)$$

### 2.3 线性二次型最优控制器设计

在这一节中, 我们着重解决多无人机系统自适应最优控制问题, 其主要内容为“领航者-跟随者”网络拓扑结构。我们通过数据驱动策略, 对无人机的输入信号、扰动信号及状态信息进行采样, 并进行一些处理操作即可解决在无人机模型未知的情况下, 多无人机编队的控制问题。这意味着我们能够实现高效的编队控制, 无论无人机内部参数的变动或模型信息的不确定性如何。

设计编队模型状态反馈分布式控制器:

$$u_i = -cK^* \left[ \sum_{j=1}^n a_{ij} (x_i - x_j - d_{ij}) + a_i (x_i - x_0 - d_i) \right] \quad (25)$$

式中:  $c$  为一个常系数;  $K^*$  为状态反馈矩阵;  $d_{ij}$  表示系统矩阵  $A$  对应的入度矩阵  $D$  中的元素, 将式 (19) 代入式 (21), 可以得到:

$$u_i = -cK^* \left\{ \sum_{j=0}^N a_{ij} \left[ (\xi_i + x_0 + d_{i0}) - (\xi_j + x_0 + d_{j0}) - d_{ij} \right] + a_i \xi_i \right\} \\ = -cK^* \left[ \sum_{j=0}^N a_{ij} (\xi_i - \xi_j) + a_i \xi_i \right] \quad (26)$$

将式 (26) 改写为编队整体:

$$u = -(L \otimes cK^* + \mathcal{L} \otimes cK^*) \zeta \quad (27)$$

式中:  $L$  为编队的拉普拉斯矩阵;  $c \geq 1/(\lambda(L + L^T + \mathcal{L} + \mathcal{L}^T))$ ;  $\mathcal{L}$  为对角矩阵, 维数与  $L$  相等; 此外对于  $i=1,2,\dots,N$ ,  $\exists i > 0$ , 使得  $a_i \neq 0$ 。

$L \otimes cK^*$  使各无人机的运动和位置状态尽量一致;  $\mathcal{L} \otimes cK^*$  使各无人机和期望状态之间的误差尽可能小, 这样能够使各个无人机都尽量跟踪领航者。

本文结合编队控制器的设计, 考虑了具体的线性二次型问题。对于此类受控系统被假设为线性系统的问题, 用控制输入状态变量和组成的二次型函

数表示系统的性能指标。通过将自适应动态规划融入线性二次型设计方法, 利用无人机的输入数据和状态数据进行迭代学习, 以近似求解最优控制器的问题。这样, 我们能够得到一个能够有效解决线性二次型最优控制问题的控制器设计方法。

对于系统来说, 线性状态反馈  $u_i = -K^* \zeta_i$  被期待用来使下面的性能指标函数成立:

$$\min_{u_i} J = \int_0^{\infty} (\zeta_i^T Q \zeta_i + u_i^T R u_i) dt \quad (28)$$

式中:  $Q = Q^T \geq 0, R = R^T \geq 0$ ,  $Q, R$  均为已知的定常矩阵, 且  $(A, \sqrt{Q})$  是可观测的。

将  $u_i = -K^* \zeta_i$  代入式 (28) 有:

$$J = \int_0^{\infty} \zeta_i^T (Q + K^{*T} R K^*) \zeta_i dt \quad (29)$$

为了求取  $K^*$ , 不妨假设存在一个常量矩阵  $P$ , 使得:

$$\frac{d}{dt} (\zeta_i^T P \zeta_i) = -\zeta_i^T (Q + K^{*T} R K^*) \zeta_i \quad (30)$$

将上式代入:

$$J = -\int_0^{\infty} \frac{d}{dt} (\zeta_i^T P \zeta_i) dt = \zeta_i^T (0) P \zeta_i (0) \quad (31)$$

将式 (30) 中的等式左侧微分形式展开:

$$\dot{\zeta}_i^T P \zeta_i + \zeta_i^T P \dot{\zeta}_i + \zeta_i^T Q \zeta_i + \zeta_i^T K^{*T} R K^* \zeta_i = 0 \\ \zeta_i^T A_c^T P \zeta_i + \zeta_i^T P A_c \zeta_i + \zeta_i^T Q \zeta_i + \zeta_i^T K^{*T} R K^* \zeta_i = 0 \quad (32) \\ \zeta_i^T (A_c^T P + P A_c + Q + K^{*T} R K^*) \zeta_i = 0$$

式中:  $A_c = A - B K^*$ 。若上式等号恒成立, 则  $A_c^T P + P A_c + Q + K^{*T} R K^*$  的值为 0。

$$(A - B K^*)^T P + P (A - B K^*) + Q + K^{*T} R K^* = 0 \quad (33) \\ A^T P + P A + Q + K^{*T} R K^* - K^{*T} B^T P - P B K^* = 0$$

令  $K^* = R^{-1} B^T P^*$ , 有:

$$A^T P^* + P^* A + Q - P^* B R^{-1} B^T P^* = 0 \quad (34)$$

式中,  $P^* = P^{*T} > 0$  (矩阵正定) 对于上述黎卡提方程来说是唯一存在的解。

公式 (34) 是一个非线性方程, 其解为矩阵  $P$ , 通过解析方法求得方程的解非常困难。

针对式 (34) 中  $P^*$  的求解问题, 本文提出一种基于数据驱动算法, 若初始状态反馈矩阵  $K_0 \in \mathbb{R}^{m \times n}$  给定, 通过求解对应  $t_k (k=0,1,\dots)$  时刻的

$P_k$  矩阵, 迭代更新线性反馈矩阵  $K_{k+1}$ , 最终近似得到黎卡提方程 (ARE) 唯一解  $P^*$ 。

**引理 1:** 在任意给定一个稳定的状态反馈矩阵  $K_0 \in \mathbb{R}^{m \times n}$  后, 我们将在  $k=0,1,2,\dots$  的情况下重复以下步骤:

1) 实对称正定矩阵  $P_k$  由下述方程求解:

$$A_k^T P_k + P_k A_k + Q + K_k^T R K_k = 0 \quad (35)$$

式中:  $A_k = A - B K_k$ 。

2) 反馈矩阵可以通过式 (36) 进行代换:

$$K_{k+1} = R^{-1} B^T P_k \quad (36)$$

有以下 3 个性质成立:

1)  $A - B K_k$  是赫尔维茨稳定的。

2)  $P^* \leq P_{k+1} \leq P_k$ 。

3)  $\lim_{k \rightarrow \infty} K_k = K^*, \lim_{k \rightarrow \infty} P_k = P^*$ 。

以上使用了策略迭代方法, 可以对策线性连续系统。最开始提供初始输入策略  $K_0$ , 该策略可以使系统稳定。然后通过式 (35) 求取  $P_k$  实现策略评估, 在式 (31) 中不难发现  $P_k$  和初始状态是与性能指标  $J$  相关联的。为了更新策略, 通过式 (36) 代换得到下一个循环中的  $K_{k+1}$ 。不过必须有较为精准的线性系统  $A, B$  的信息来实现引理 1 中的迭代循环过程, 本文已给出 2 个与其相关的矩阵。若无人机系统存在参数摄动, 则精确的矩阵信息  $\hat{A}, \hat{B}$  将无法获得, 所以引理 1 所提出的策略迭代方法实用性较差。

在这篇论文中, 结合了引理 1 和自适应动态规划的思想, 基于数据驱动, 我们提出一种自适应最优控制算法。这种方法在适用于明确已知的线性系统模型的同时, 只需要一段时间内的输入和状态信息, 即可求解出最优的反馈控制策略。这意味着我们可以通过数据驱动的方法来实现无人机的最优控制, 无需依赖于准确的系统模型。

将第  $i$  架无人机与期望位置之间的误差变量的数学模型改写如下:

$$\dot{\zeta}_i = \hat{A}_k \zeta_i + \hat{B}(K_k \zeta_i + u_i) \quad (37)$$

式中:  $\hat{A}_k = \hat{A} - \hat{B} K_k$ 。若将性能指标函数在  $[t, t + \Delta t]$  积分, 有:

$$\begin{aligned} \zeta_i^T P_k \zeta_i \Big|_t^{t+\Delta t} = & \int_t^{t+\Delta t} -\zeta_i^T \left( Q + (K_k)^T R K_k \right) \zeta_i d\tau + \\ & 2 \int_t^{t+\Delta t} (u_i + K_k \zeta_i)^T R K_{k+1} \zeta_i d\tau \end{aligned} \quad (38)$$

将上式中的系统参数代入可转移性能指标函数, 以克隆克积的形式表示内部多项式如下:

$$\begin{aligned} \zeta_i^T \left( Q + (K_k)^T R K_k \right) \zeta_i = & \left( \zeta_i^T \otimes \zeta_i^T \right) \text{vec} \left( Q + (K_k)^T R K_k \right) \\ & \left( u_i + K_k \zeta_i \right)^T R K_{k+1} \zeta_i = \\ & \left[ \left( \zeta_i^T \otimes \zeta_i^T \right) \left( I_n \otimes (K_k)^T R \right) + \right. \\ & \left. \left( \zeta_i^T \otimes u_i^T \right) \left( I_n \otimes R \right) \right] \text{vec} \left( K_{k+1} \right) \end{aligned} \quad (39)$$

式中,  $\text{vec} \left( Q + (K_k)^T R K_k \right)$  表示矩阵  $Q + (K_k)^T R K_k$  的向量形式, 在迭代求解 (39)  $K_{k+1}$  时, 本文使用了最小二乘法的方式进行。给定正整数  $s, n_\mu, n_v, n_w$ ; 定义下面的向量,  $\hat{\mu} \in \mathbb{R}^{(1/2)n_\mu(n_\mu+1)}, \Gamma_{v,w} \in \mathbb{R}^{s \times n_v n_w}, \xi_{v,v} \in \mathbb{R}^{s \times (1/2)n_v(n_v+1)}$ :

$$\begin{aligned} \hat{\mu} = & [\mu_1^2, \mu_1 \mu_2, \dots, \mu_1 \mu_\mu^n, \mu_2^2, \mu_2 \mu_3, \dots, \mu_\mu^n - 1, \mu_\mu^n, \mu_\mu^2]^T \\ \Gamma_{v,w}^i = & \left[ \int_{t_{i,0}}^{t_{i,1}} v \otimes w d\tau, \int_{t_{i,1}}^{t_{i,2}} v \otimes w d\tau, \dots, \int_{t_{i,s-1}}^{t_{i,s}} v \otimes w d\tau \right]^T \\ \xi_{\hat{\mu}, \hat{\mu}}^i = & [\hat{\mu}(t_i, 1) - \hat{\mu}(t_i, 0), \hat{\mu}(t_i, 2) - \hat{\mu}(t_i, 1), \dots, \hat{\mu}(t_i, s) - \\ & \hat{\mu}(t_i, s-1)]^T \\ \hat{Q} = & [2\Omega_{11}, 2\Omega_{12}, \dots, 2\Omega_{1n}, \Omega_{22}, 2\Omega_{23}, \dots, \Omega_{n-1,n}, \Omega_{nn}]^T \end{aligned}$$

式中:  $0 \leq t_{i,0} < t_{i,1} < \dots < t_{i,s} < t_{i+1,0} < t_{i+1,1} < \dots$ ,  $s$  表示在  $t_i$  与  $t_{i+1}$  之间的采样次数。

将式 (39) 扩展成如下形式:

$$\Theta_k \begin{bmatrix} \hat{P}_k \\ \text{vec} \left( K_{k+1} \right) \end{bmatrix} = \Phi_k \quad (40)$$

式中:

$$\begin{aligned} \Theta_k = & \left[ \xi_{\hat{\mu}, \hat{\mu}}^k, -2\Gamma_{\zeta_i, \zeta_i}^k \left( I_n \otimes (K_k)^T R \right) - 2\Gamma_{\zeta_i, u_i}^k \left( I_n \otimes R \right) \right] \\ \Phi_k = & -\Gamma_{\zeta_i, \zeta_i}^k \text{vec} \left( Q + (K_k)^T R K_k \right) \end{aligned}$$

在利用最小二乘法求解式 (40)  $K_{k+1}$  时, 解的唯一性是非常关键的因素, 而满秩是确定解唯一性的充分条件, 即  $\Theta_k$  必须做到满秩。使  $\Theta_k$  满秩的充分条件如下。

**引理 2:** 对于每次迭代  $k=0,1,\dots$ ，存在一个充分大的  $l_k > 0$ ，如果有下面的秩条件成立：

$$\text{rank}\left([\Gamma_{\zeta_i, \zeta_i}^k, \Gamma_{\zeta_i, u_i}^k]\right) = \frac{n(n+1)}{2} + mn \quad (41)$$

那么式 (40) 中  $\Theta_k$  满秩。为了节省空间，此处省略证明。

式 (41) 成立时，式 (40) 有唯一解，应用最小二乘法求解，式 (40) 进一步写成：

$$\begin{bmatrix} \hat{P}_k \\ \text{vec}(\mathbf{K}_{k+1}) \end{bmatrix} = ((\Theta_k)^T \Theta_k)^{-1} (\Theta_k)^T \Phi_k \quad (42)$$

假设每个区间  $[t_{i,0}, t_{i,s}]$  为一个采样周期，在解决上述方程时，必须有充足的样本数据，即迭代过程中，都应该存在充分大的  $s$  进行预处理以确保数据的可靠。首先，考虑到零输入时本文解耦得到的子系统具有发散性，为使系统保持稳定，我们可以使用初始控制策略  $\mathbf{K}_0$ ；其次，选择探测噪声也是非常重要的。一般而言，计算秩条件可以检查最小二乘法中的秩问题，但无法通过分析进行确认。

为了满足秩条件，应确保每个迭代步骤至少有 10 个相对应的采样周期的数据，即  $s \geq 10$ 。当探测噪声有周期性时，其周期应远小于采样时间间隔。

由控制器设计  $u_i = -\mathbf{K}^* \zeta_i$ 、式 (40) 以及式 (50) 构成基于数据驱动的自适应最优控制算法，流程图如图 1 所示。其中  $c$  为常数，表示  $P_k$  需要达到的误差精度， $\omega$  表示控制输入的参数摄动。

对于具有“领航者-跟随者”有向通信拓扑结构的多无人机系统，上述基于数据驱动的自适应最优控制器能够有效地解决在系统内部参数摄动情况下的编队控制器设计问题。接下来，我们将提供该控制算法的可行性证明，以验证该算法的有效性和可行性。

**定理 1:** 若有初始控制策略  $\mathbf{K}_0 \in \mathbb{R}^{m \times n}$  使系统稳定的同时满足引理 2 中的秩条件，那么方程式的解迭代序列  $\{P_k\}_{k=0}^{\infty}, \{K_k\}_{k=1}^{\infty}$  可以收敛到相应的  $P^*, K^*$ 。

首先证明该策略迭代方法的收敛性。

由式 (30) 可得：

$$P = \int_0^{\infty} e^{(A-BK)^T t} (Q + K^T R K) e^{(A-BK)t} dt \quad (43)$$

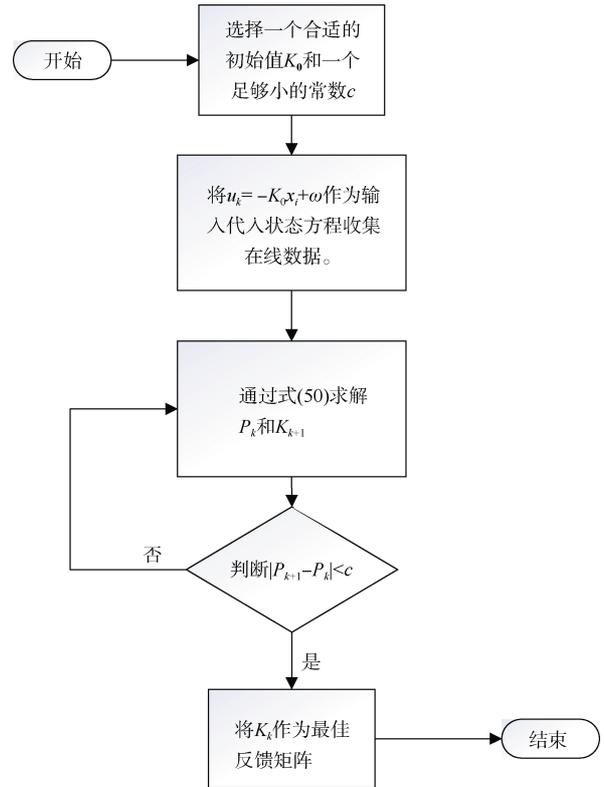


图 1 基于数据驱动的自适应最优控制算法流程图  
Fig. 1 Flowchart of data-driven adaptive optimal control algorithm

可以推出：

$$P_k - P_{k+1} = \int_0^{\infty} e^{(A-BK_{k+1})^T t} [(K_k - K_{k+1})^T R (K_k - K_{k+1}) - (K_k - K_{k+1})^T (B^T P_k - R L_{k+1}) - (B^T P_k - R L_{k+1})^T (K_k - K_{k+1})] e^{(A-BK_{k+1})t} dt \quad (44)$$

或者：

$$P_k - P_{k+1} = \int_0^{\infty} e^{(A-BK_k)^T t} [(K_k - K_{k+1})^T R (K_k - K_{k+1}) - (K_k - K_{k+1})^T (B^T P_{k+1} - R L_{k+1}) - (B^T P_{k+1} - R L_{k+1})^T (K_k - K_{k+1})] e^{(A-BK_k)t} dt \quad (45)$$

当  $k=0$  时， $K_1 = R^{-1} B P_0$ ，代入上式：

$$P_0 - P_1 = \int_0^{\infty} e^{(A-BK_1)^T t} [(K_0 - K_1)^T R (K_0 - K_1)] e^{(A-BK_1)t} dt \geq 0 \quad (46)$$

根据式 (45) 有：

$$P_1 - P^* = \int_0^{\infty} e^{(A-BK_1)^T t} [(K_1 - K^*)^T R (K_1 - K^*)] e^{(A-BK_1)t} dt \geq 0 \quad (47)$$

因此,  $P^* \leq P_1 \leq P_0$ 。又因为  $P^*$  是黎卡提方程的解, 必为正定矩阵,  $P_0$  正定且有界, 所以  $A - BK_1$  是符合 Hurwitz 稳定性判据的。令  $k=0, 1, \dots$ , 通过上式反复进行验证, 可得  $P^* \leq P_{k+1} \leq P_k$ 。由于  $P_k$  为递减的矩阵序列且有下界  $P^*$ , 故  $\lim_{k \rightarrow \infty} P_k = P_\infty$  存在,  $P_\infty$  满足式 (34) 且是该方程的唯一解。因此有  $\lim_{k \rightarrow \infty} K_k = K^*$ ,  $\lim_{k \rightarrow \infty} P_k = P^*$ , 该迭代策略是收敛的。

当给定一个使系统稳定的状态反馈矩阵  $K_k$ , 如果  $P_k$  是式 (33) 的解, 并且  $K_{k+1}$  通过  $K_{k+1} = R^{-1}B^T P_k$  唯一决定。由式 (38) 可得,  $K_k$ 、 $P_k$  满足引理 1 中提到的 3 个性质。另一方面, 如果令  $P_k = P$ , 且  $K$  使下式成立:

$$\Theta_k \begin{bmatrix} \hat{P} \\ \text{vec}(K) \end{bmatrix} = \Phi_k \quad (48)$$

可以得出,  $\hat{P} = \hat{P}_k, \text{vec}(K) = \text{vec}(K_{k+1})$ 。由引理 2 可知  $P, K$  是唯一的, 所以  $\hat{P}, \text{vec}(K)$  也具有唯一性。

### 3 稳定性分析

**定理 2:** 对于公式 (21), 当  $P^*, K^*$  为 3.3 节中求出的线性二次型最优反馈矩阵时, 线性反馈控制器  $u = -K^* \zeta$  可以使编队稳定, 且使无人机编队系统性能指标能够达到最小。

针对公式 (21) 选取李雅普诺夫函数:

$$V(t) = \frac{1}{2} \zeta^T(t) (I_N \otimes P^*) \zeta(t) \quad (49)$$

对等式两边进行求导, 并代入式 (21) 可得:

$$\begin{aligned} \dot{V} &= \zeta^T (I_N \otimes P^*) \dot{\zeta} \\ &= \frac{1}{2} \zeta^T [I_N \otimes (P^* \hat{A} + \hat{A}^T P^*) - (L + L^T) \otimes cP^* \hat{B}R^{-1} \hat{B}^T P^* - \\ &\quad (L + L^T) \otimes cP^* \hat{B}R^{-1} \hat{B}^T P^*] \zeta \end{aligned} \quad (50)$$

由于  $c \geq 1/(\lambda(L + L^T + L + L^T))$ , 则可得:

$$c(L + L^T + L + L^T) \geq I_N \quad (51)$$

得到:

$$\begin{aligned} \dot{V} &\leq \frac{1}{2} \zeta^T [I_N \otimes (P^* \hat{A} + \hat{A}^T P^* - P^* \hat{B}R^{-1} \hat{B}^T P^*)] \zeta \\ &= -\frac{1}{2} \zeta^T [I_N \otimes Q] \zeta \leq 0 \end{aligned} \quad (52)$$

由此可得, 所选的 Lyapunov 函数  $V(t) > 0$  且  $\dot{V}(t) \leq 0$ 。本文提出的编队控制器可以使编队状态稳定, 即

$$\lim_{t \rightarrow \infty} \zeta = 0 \Rightarrow \lim_{t \rightarrow \infty} \zeta_i = x_i - \varrho_0 - \varrho_{i0} = 0 \quad (53)$$

在上述控制策略下的无人机编队系统性能指标为

$$J = \sum_{i=1}^N \int_0^\infty (\zeta_i^T Q \zeta_i + u_i^T R u_i) dt \quad (54)$$

又因为  $P^*$  是黎卡提方程 (ARE) 的唯一解,  $K^* = R^{-1}B^T P^*$ , 由 LQR 的原理可知,  $u = -K^* \zeta$  为最优控制器, 可使上述系统性能指标达到最小。

### 4 仿真分析

为了验证本文提出的控制方法的有效性, 在本节中, 我们进行了基于数据驱动的自适应最优控制编队的仿真实验。实验假设无人机可以对状态数据进行交换, 并确保编队网络中至少存在一个有向连通图。

在本实验中, 我们验证了一个由 4 架四旋翼无人机构成的编队。该系统的通信拓扑图如图 2 所示。

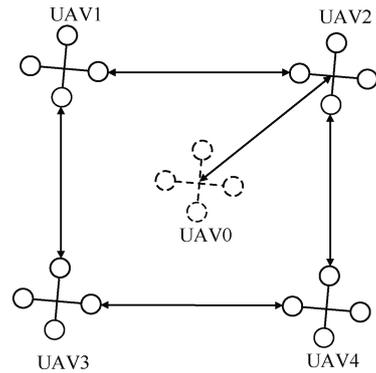


图 2 编队通信拓扑图

Fig. 2 Formation communication topology diagram

图中: UAV0 表示虚拟领航者, UAV1~UAV4 表示跟随者无人机。假定只有 UAV2 与 UAV0 有信息交换。

图 2 所对应的入度矩阵与邻接矩阵为

$$A = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}, \quad D = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

由此得到该编队系统的拉普拉斯矩阵为

$$L = D - A = \begin{bmatrix} 2 & -1 & -1 & 0 \\ -1 & 2 & 0 & -1 \\ -1 & 0 & 2 & -1 \\ 0 & -1 & -1 & 2 \end{bmatrix}$$

由信息交换能够进行与否，可以判断  $a_i = 0 (i=1,3,4), a_2 = 2$ 。其中  $a_2$  的数值决定了编队状态趋近于期望状态的速度，故应在允许范围内选择尽量大的值。

假设编队队形是一个平面正方形，其中领航者位于正方形的几何中心。

在上述假设下队形矩阵如下描述：

$$d_{ij}^x = \begin{bmatrix} 0 & 2 & 0 & 2 \\ -2 & 0 & -2 & 0 \\ 0 & 2 & 0 & 2 \\ -2 & 0 & -2 & 0 \end{bmatrix}, d_{ij}^y = \begin{bmatrix} 0 & 0 & -2 & -2 \\ 0 & 0 & -2 & -2 \\ 2 & 2 & 0 & 0 \\ 2 & 2 & 0 & 0 \end{bmatrix},$$

$$d_{ij}^z = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

跟随者与领航者的距离矩阵为

$$d_i^x = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}, d_i^y = \begin{bmatrix} -1 \\ 1 \\ -1 \\ 1 \end{bmatrix}, d_i^z = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

各个跟随者相对于虚拟领航者的的初始状态信息。

表 1 各无人机初始状态  
Table 1 Initial state of each UAV

| 无人机编号 | 相对位置 / (m, m, m) | 速度 / (m·s <sup>-1</sup> ) | 偏航角 / (°) |
|-------|------------------|---------------------------|-----------|
| UAV1  | (8, 6, -2)       | (-3, -1, 6)               | $\pi/6$   |
| UAV2  | (10, 1, 1.5)     | (5, -7, -1)               | $\pi/7$   |
| UAV3  | (3, -8, 3)       | (-5, 3, 4)                | $\pi/3$   |
| UAV4  | (-5, -5, -4)     | (4, 4, -2)                | $\pi/4$   |

首先，求解位置子系统  $x$  方向，假设在求解时  $Q=I_4, R=1$ 。无人机的初始状态数据为  $x_i = [5, 2, 2, 1]^T$ 。采样周期设定为  $T=0.01$  s，采样时间设定为 5 s，初始的状态反馈矩阵为  $K_0 = [1, 3, 4, 3]$ ，探测噪声由周期信号叠加：

$$0.2 \times \sum_i \sin(i \times t), (i=1, 3, 7, 11, 13, 15)$$

仿真结果如下：

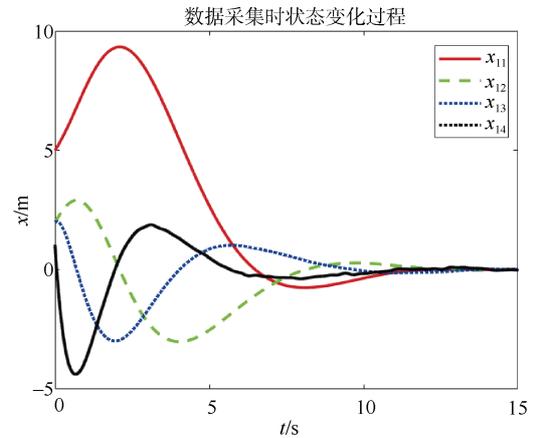


图 3 位置  $x$  数据采集时的状态响应曲线  
Fig. 3 Status response curve for data acquisition of position  $x$

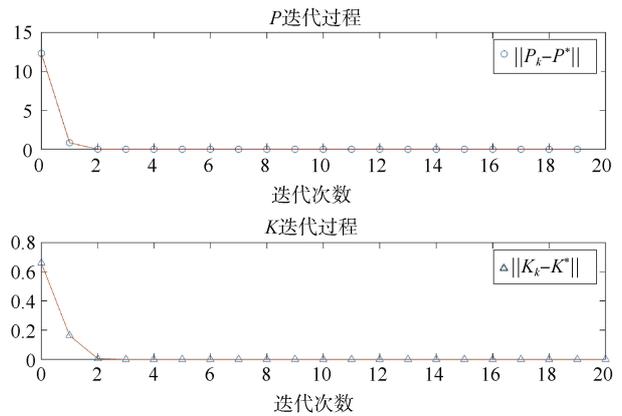


图 4 位置  $x$  子系统最优反馈增益迭代过程  
Fig. 4 Optimal feedback gain iteration process of position  $x$  subsystem

从图 3 可以观察到，在给定初始状态条件下，无人机最终趋于稳定。因此，整个过程收集到的数据均有效，并满足引理 2 对数据矩阵秩的要求。图 5 具体表现了反馈增益矩阵的迭代过程，即每次迭代得到的反馈增益矩阵与期望的差值。三次迭代后算法基本收敛，这成功地验证了定理 1。得出最优反馈增益矩阵如下：

$$P^* = \begin{bmatrix} 3.0777 & 4.2361 & 3.0777 & 1 \\ 4.2361 & 9.9596 & 8.4721 & 3.0777 \\ 3.0777 & 8.4721 & 9.9596 & 4.2361 \\ 1 & 3.0777 & 4.2361 & 3.0777 \end{bmatrix}$$

$$K^* = [1 \quad 3.0777 \quad 4.2361 \quad 3.0777]$$

下面给出编队的状态响应曲线：

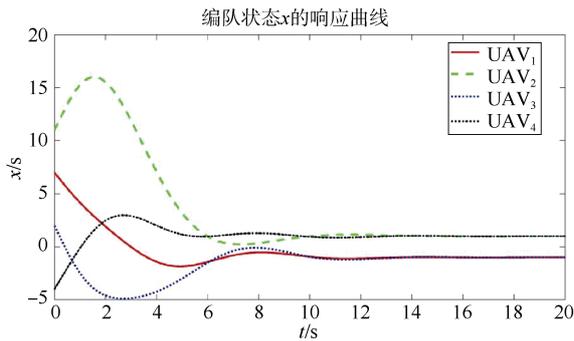


图 5 编队位置状态  $x$  的误差变量响应曲线  
Fig. 5 Error variable response curve for formation position state  $x$

从图 5 可以看出, 编队能在 15 s 内收敛到期望状态。

由对称性可知, 状态子系统  $x$  与  $y$  是一致的, 受篇幅所限, 不在此给出位置状态  $y$  的仿真结果。

对于位置状态  $z$  的控制器, 其简化后是一个二阶系统, 与  $x, y$  均不同, 所以必要获得的未知参数变少。

假设该最优状态控制器的  $Q=I_2, R=1$ 。在线数据采集时无人机的初始状态数据为  $z_i = [3, -2]^T$ 。数据采集的步骤中, 采样周期和采样时间分别设置为 0.01 s、2 s, 迭代开始的状态反馈矩阵为  $K_0 = [1, 1.5]$ , 探测噪声同样选择周期信号的叠加:

$$0.2 \times \sum_i \sin(i \times t), (i = 1, 3, 7, 11, 13, 15)$$

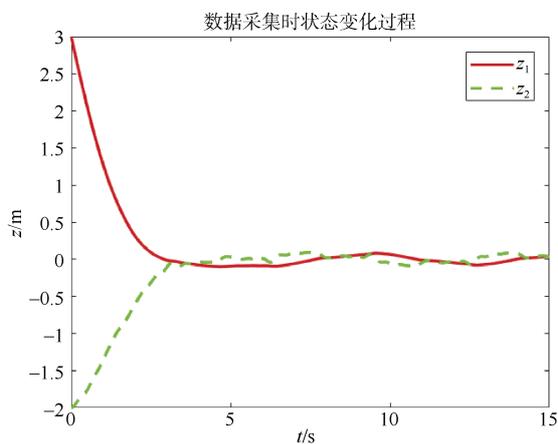


图 6 位置  $z$  数据采集时的状态响应曲线  
Fig. 6 Status response curve for data acquisition of position  $z$

图 7 展示的是状态  $z$  控制器的迭代过程。因为

系统的阶数相对较低, 所以状态的收敛速度更快, 第二次迭代后, 得到了最优状态增益矩阵:

$$P^* = \begin{bmatrix} 1.7321 & 1 \\ 1 & 1.7321 \end{bmatrix}$$

$$K^* = [1 \quad 1.7321]$$

观察图 8 后发现, 在  $z$  方向只需 8 s 的调整即可使状态达到稳定。

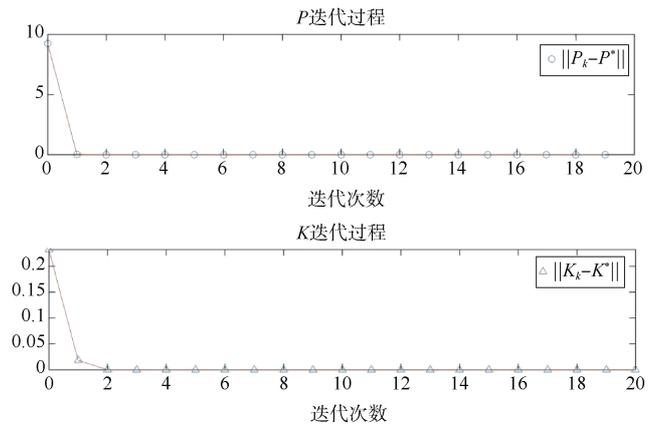


图 7 位置  $z$  子系统最优反馈增益迭代过程  
Fig. 7 Optimal feedback gain iteration process of position  $z$  subsystem

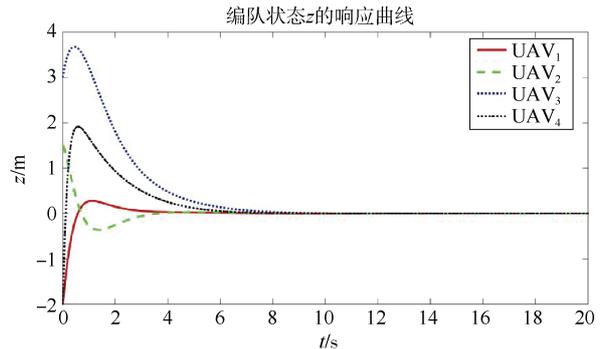


图 8 状态  $z$  的误差变量响应曲线  
Fig. 8 Error variable response curve for state  $z$

当虚拟领航者的期望状态如下:

$$\rho_0^x = \begin{bmatrix} 5 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \rho_0^y = \begin{bmatrix} 5 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \rho_0^z = \begin{bmatrix} 5 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

图 9 显示了编队运动的仿真结果, 图像化了编队实现悬停的调节过程。因为简化后偏航角子系统的模型与状态  $z$  子系统相似, 受篇幅所限, 不在此给出编队姿态角的误差仿真结果。

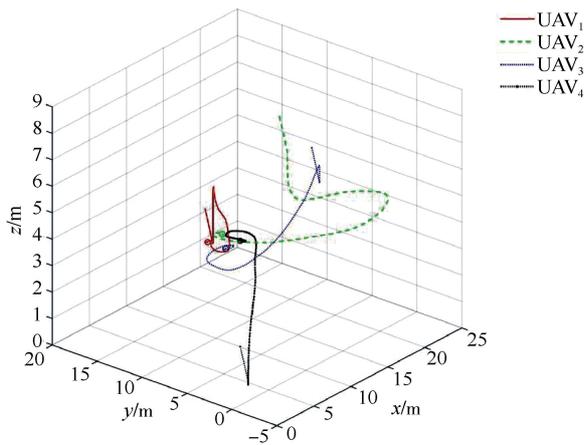


图9 编队悬停三维飞行轨迹

Fig. 9 Formation hovers over a three-dimensional flight trajectory

## 5 结束语

本文研究了在存在系统参数摄动情况下的四旋翼无人机编队问题,考虑到无人机编队在实际应用中的广阔前景以及对复杂任务需求的完成。通过结合分布式一致性和线性二次型理论,为了解决无人机参数摄动导致的模型参数未知问题,利用数据驱动构造了自适应动态规划的最优编队控制器的方法。该方法对每个子系统进行编队控制设计,从而采用了虚拟领航-跟随者模型,并简化了四旋翼无人机动力学的非线性系统,将其简化为4个线性子系统,随后提供了详细的设计步骤和证明。最后,通过 Lyapunov 方法分析论证了算法的稳定性,并给出了仿真验证的结果。这为以后在真实环境中实现四旋翼无人机编队飞行提供了坚实的理论支持。

## 参考文献

- [1] ZONG Q, WANG D, SHAO S K, et al. Research status and development of multi-UAV collaborative formation flight control. *Journal of Harbin Institute of Technology*[J]. *Journal of Harbin Institute of Technology*, 2017, 49 (3): 1-14.
- [2] 庄亚楠, 苗国英, 孙英博. 基于事件触发的多无人机群快速编队控制[J]. *电子信息对抗技术*, 2022, 37 (6): 48-53.
- [3] JI L, FAN C X, QIAO Y, et al. Formation control for multi-UAVs system based on event-triggered mechanism[C]// 2019 Chinese Control Conference (CCC). Guangzhou: IEEE, 2019.
- [4] KHACHUMOV M. An approach to formation control of UAVs based on applying adapted Kohonen neural network[C]// 2023 IEEE Ural-Siberian Conference on Biomedical Engineering, Radioelectronics and Information Technology ( USBEREIT ). Yekaterinburg: IEEE, 2023.
- [5] KADA B, KHALID M, SHAIKH M S. Distributed cooperative control of autonomous multi-agent UAV systems using smooth control[J]. *Journal of Systems Engineering and Electronics*, 2020, 31 (6): 1297-1307.
- [6] GHAMRY K A, DONG Y Q, KAMEL M A, et al. Real-time autonomous take-off, tracking and landing of UAV on a moving UGV platform[C]// 24th Mediterranean Conference on Control and Automation ( MED ). Athens: IEEE, 2016.
- [7] ZHANG J L, YAN J G, ZHANG P. Multi-UAV formation control based on a novel back-stepping approach[J]. *IEEE Transactions on Vehicular Technology*, 2020, 69 (3): 2437-2448.
- [8] GU Z H, SONG B J, FAN Y S, et al. Design and verification of UAV formation controller based on leader-follower method[C]// 7th International Conference on Automation, Control and Robotics Engineering ( CACRE ). Xi'an: IEEE, 2022.
- [9] ULLAH N, MEHMOOD Y, ASLAM J, et al. UAVs-UGV leader follower formation using Adaptive non-singular terminal super twisting sliding mode control[J]. *IEEE Access*, 2021, 9: 74385-74405.
- [10] HOU Z S, XU J X. On data-driven control theory: the state of the art and perspective[J]. *Acta Automatica Sinica*, 2009, 35: 650-667.
- [11] HELVOORT J J M. Unfalsified control: data-driven control design for performance improvement[D]. Eindhoven: Technische Universiteit Eindhoven, 2007.
- [12] HEUSDEN K. Non-iterative data-driven model reference control[D]. Lausanne: EPFL Lausanne, 2010.
- [13] HOU Z S, XU J X. Review and prospect of data-driven control theory and methodology[J]. *Journal of Automatica Sinica*, 2009, 35 (6): 650-667.
- [14] KLEINMAN D L. On an iterative technique for Riccati equation computations[J]. *IEEE Transactions on Automatic Control*, 1968, 13 (1): 114-115.

(责任编辑:肖楚楚)