



(12) 发明专利

(10) 授权公告号 CN 113268074 B

(45) 授权公告日 2022.05.13

(21) 申请号 202110632549.2

CN 103017753 A, 2013.04.03

(22) 申请日 2021.06.07

CN 106595671 A, 2017.04.26

(65) 同一申请的已公布的文献号

JP 2007064691 A, 2007.03.15

申请公布号 CN 113268074 A

WO 2021017227 A1, 2021.02.04

US 10671076 B1, 2020.06.02

(43) 申请公布日 2021.08.17

丁力等. 基于航迹规划的四旋翼飞行器轨迹跟踪控制.《电光与控制》.2018, (第11期), 34-39.

(73) 专利权人 哈尔滨工程大学

地址 150001 黑龙江省哈尔滨市南岗区南通大街145号哈尔滨工程大学科技处知识产权办公室

万磊, 等. 欠驱动水下机器人航迹跟踪控制.《电机与控制学报》.2013, 第17卷(第2期), 103-111.

(72) 发明人 高敬鹏 胡欣瑜 叶方 江志焯

毛新蕊 高路 郑沛 何重航

Jiaxuan Fan等.UAV online path planning technology based on deep reinforcement learning.《2020 Chinese Automation Congress (CAC)》.2020, 5382-5386.

(51) Int. Cl.

G05D 1/08 (2006.01)

G05D 1/10 (2006.01)

宋欣屿, 等. 基于深度强化学习的无人机着陆轨迹跟踪控制.《航空科学技术》.2020, 第31卷(第1期), 68-75.

(56) 对比文件

CN 110162077 A, 2019.08.23

CN 110333739 A, 2019.10.15

CN 103471592 A, 2013.12.25

CN 112819253 A, 2021.05.18

审查员 王玮

权利要求书3页 说明书12页 附图3页

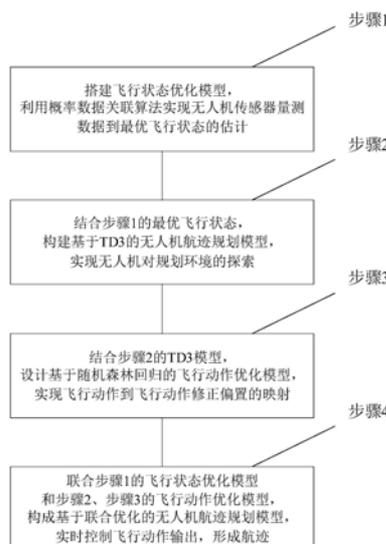
(54) 发明名称

一种基于联合优化的无人机航迹规划方法

航迹规划。

(57) 摘要

本发明属于无人机飞行控制技术领域,具体涉及一种基于联合优化的无人机航迹规划方法。本发明通过前端飞行状态优化模型获取无人机自身飞行状态的最优估计,大幅度降低了自然干扰因素对传感器数据的误差;通过后端飞行动作优化模型,将实际偏差大的飞行角度调整至预期飞行角度,解决了强化学习决策飞行动作不佳的问题,提高了飞行航迹的可靠性。在实际应用中,对于传感器实时采集飞行状态数据,利用前端飞行状态优化模型实时获得最优飞行状态估计,将其作为已训练TD3模型的输入,得到飞行动作,并利用后端飞行动作优化模型,得到对应的飞行动作偏置,对其进行判决,控制输出优化后的飞行动作,从而实现无人机面对自然干扰影响的实时



CN 113268074 B

1. 一种基于联合优化的无人机航迹规划方法,其特征在于,包括以下步骤:

步骤1:搭建飞行状态优化模型;所述的飞行状态优化模型的输入为无人机携带的I个传感器在t时刻对无人机飞行状态的量测向量 $m_i(t)$,飞行状态优化模型的输出为无人机t时刻飞行状态的最优估计 $\zeta(t|t)=[p(t),v(t)]^T$;

步骤1.1:输入无人机携带的I个传感器在t时刻对无人机飞行状态的量测向量 $m_i(t)$;

其中,I个传感器的采样间隔均为 Ω ; $m_i(t)=[p_i(t),v_i(t)]^T$; $i=\{1,2,\dots,I\}$; $p_i(t)$ 表示第i个传感器获取的无人机位置量测向量, $p_i(t)=[p_{ix}(t),p_{iy}(t),p_{iz}(t)]^T$, $p_{ix}(t)$ 、 $p_{iy}(t)$ 和 $p_{iz}(t)$ 分别表示第i个传感器获取的笛卡尔坐标系下无人机在三维空间中x、y和z三个方向的位置坐标量测值; $v_i(t)$ 表示第i个传感器获取的无人机速度量测向量, $v_i(t)=[v_{ix}(t),v_{iy}(t),v_{iz}(t)]^T$, $v_{ix}(t)$ 、 $v_{iy}(t)$ 和 $v_{iz}(t)$ 分别表示第i个传感器获取的笛卡尔坐标系无人机在三维空间中x、y和z三个方向的速度分量量测值;

步骤1.2:根据t-1时刻无人机飞行状态的最优估计 $\zeta(t-1|t-1)$,计算t时刻无人机飞行状态的预测 $\zeta(t|t-1)$;

$$\zeta(t|t-1)=\mathbf{F}\zeta(t-1|t-1)+\tilde{\omega}(t)$$

其中, $\zeta(t-1|t-1)=[p(t-1),v(t-1)]^T$; $\mathbf{F}=\begin{bmatrix} 1 & \Omega \\ 1 & 1 \end{bmatrix}$; $\tilde{\omega}(t)$ 为白噪声;

步骤1.3:计算t时刻各传感器对无人机飞行状态的量测预测 $\delta_i(t|t-1)$;

$$\delta_i(t|t-1)=\mathbf{H}_i(t)\zeta(t|t-1)$$

其中, $\mathbf{H}_i(t)$ 表示第i个传感器的线性测量矩阵; $\mathbf{H}_i(t)=\begin{bmatrix} \frac{p_i(t)}{\|p_i(t)\|} & 1 \\ 1 & \frac{v_i(t)}{\|v_i(t)\|} \end{bmatrix}$; $\|\cdot\|$ 表示求

模运算;

步骤1.4:计算t时刻各传感器的量测新息 $\varepsilon_i(t)$;

$$\varepsilon_i(t)=m_i(t)-\delta_i(t|t-1)$$

步骤1.5:根据t-1时刻无人机飞行状态的协方差 $P(t-1|t-1)$,计算t时刻无人机飞行状态的协方差预测 $P(t|t-1)$;

$$P(t|t-1)=\mathbf{F}P(t-1|t-1)\mathbf{F}^T+\mathbf{O}$$

$$\text{其中, } \mathbf{O}=\begin{bmatrix} \frac{\Omega^3}{3} & \frac{\Omega^2}{2} \\ \frac{\Omega^2}{2} & \Omega \end{bmatrix};$$

步骤1.6:计算t时刻各传感器的预测新息协方差 $\mathbf{I}_i(t)$;

$$\mathbf{I}_i(t)=\mathbf{H}_i(t)P(t|t-1)\mathbf{H}_i^T(t)+\mathbf{R}_i(t)$$

$$\text{其中, } \mathbf{R}_i(t)=\begin{bmatrix} 0.001 * m_i(t) & 0 \\ 0 & \frac{1}{m_i(t)} \end{bmatrix};$$

步骤1.7: 计算t时刻各传感器量测向量的关联概率 $\beta_i(t)$;

$$\beta_i(t) = \frac{e_i}{\sum_{i=1}^I e_i}$$

$$e_i = \frac{1}{2\pi \|\mathbf{I}_i(t)\|^{\frac{1}{2}}} \exp\left[-\frac{1}{2} \boldsymbol{\varepsilon}_i^T(t) \mathbf{I}_i^{-1}(t) \boldsymbol{\varepsilon}_i(t)\right]$$

步骤1.8: 计算t时刻I个传感器的综合量测新息 $\boldsymbol{\varepsilon}(t)$;

$$\boldsymbol{\varepsilon}(t) = \sum_{i=1}^I \beta_i(t) \boldsymbol{\varepsilon}_i(t)$$

步骤1.9: 计算t时刻无人机飞行状态的协方差 $\mathbf{P}(k|k)$;

$$\mathbf{P}(t|t) = \mathbf{K}(t) \left[\sum_{i=1}^I \beta_i(t) \boldsymbol{\varepsilon}_i(t) \boldsymbol{\varepsilon}_i^T(t) - \boldsymbol{\varepsilon}(t) \boldsymbol{\varepsilon}^T(t) \right] \mathbf{K}^T(t)$$

$$\mathbf{K}(t) = \sum_{i=1}^I \mathbf{P}(t|t-1) \mathbf{H}_i^T(t) \mathbf{I}_i^{-1}(t)$$

步骤1.10: 计算无人机t时刻飞行状态的最优估计 $\zeta(t|t)$;

$$\zeta(t|t) = \zeta(t|t-1) \mathbf{K}(t) \boldsymbol{\varepsilon}(t)$$

步骤2: 构建并训练基于TD3的无人机航迹规划模型;

设置基于TD3的无人机航迹规划模型的强化学习状态空间为 $\mathbf{s}_t = [\mathbf{p}_t, \mathbf{v}_t]^T$; $\mathbf{p}_t = [\mathbf{p}_{x,t}, \mathbf{p}_{y,t}, \mathbf{p}_{z,t}]^T$ 表示t时刻无人机位置信息; $\mathbf{p}_{x,t}$ 、 $\mathbf{p}_{y,t}$ 和 $\mathbf{p}_{z,t}$ 分别代表笛卡尔坐标系下, 无人机在三维空间中x、y和z三个方向的位置坐标点; $\mathbf{v}_t = [\mathbf{v}_{x,t}, \mathbf{v}_{y,t}, \mathbf{v}_{z,t}]^T$ 为t时刻无人机速度信息, $\mathbf{v}_{x,t}$ 、 $\mathbf{v}_{y,t}$ 和 $\mathbf{v}_{z,t}$ 分别代表笛卡尔坐标系下, 无人机在三维空间中x、y和z三个方向的速度分量;

设置基于TD3的无人机航迹规划模型的状态空间输入为 $\zeta(t|t)$, 设置基于TD3的无人机航迹规划模型的动作空间的输出为 $\mathbf{a}(t) = [\boldsymbol{\varphi}(t), \boldsymbol{\theta}(t)]$, $\boldsymbol{\varphi}(t)$ 表示无人机t时刻飞行的方位角, $\boldsymbol{\theta}(t)$ 表示无人机t时刻飞行的俯仰角; 设置强化学习奖励函数 $r(t)$ 为:

$$r(t) = r_1(t) + r_2(t)$$

$$r_1(t) = \begin{cases} 1 - \|\mathbf{p}(t) - \mathbf{p}_{arrive}\| / 25 & \|\mathbf{p}(t) - \mathbf{p}_{arrive}\| < d_{max} \\ -\|\mathbf{p}(t) - \mathbf{p}_{arrive}\| / 25 & \|\mathbf{p}(t) - \mathbf{p}_{arrive}\| \geq d_{max} \end{cases}$$

$$r_2(t) = \begin{cases} \|\mathbf{p}(t) - \mathbf{p}_{start}\| / 25 & \|\mathbf{p}(t) - \mathbf{p}_{start}\| < \rho_{max} \\ -2 & \|\mathbf{p}(t) - \mathbf{p}_{start}\| \geq \rho_{max} \end{cases}$$

其中, r_1 表示到达正奖励; r_2 表示航程负奖励; \mathbf{p}_{arrive} 表示终点的位置坐标; \mathbf{p}_{start} 表示起点的位置坐标; d_{max} 表示无人机最大探测范围; ρ_{max} 表示无人机携带的燃料可供的最大可飞行航程;

搭建基于TD3的无人机航迹规划模型, 其中Actor网络和Critic网络均为3层全连接网络结构, Actor网络的输入层为2个神经元, 输出层为6个神经元, Critic网络的输入层为8个神经元, 输出层为1个神经元, 设置训练参数, 进行网络训练, 得到基于TD3的无人机航迹规划模型;

步骤3: 构建并训练基于随机森林回归的飞行动作优化模型;

利用步骤2中训练好的基于TD3的无人机航迹规划模型, 根据每一时刻基于TD3的无人机航迹规划模型输出的动作 $a(t) = [\varphi(t), \vartheta(t)]$, 获取无人机执行动作 $a(t)$ 后在 $t+1$ 时刻到达的位置, 将无人机 $t+1$ 时刻的位置与 t 时刻的位置连接成标定线, 标定线与无人机执行动作 $a(t)$ 所得航迹构成夹角 $\tilde{a}(t)$, 在直角坐标系中 $\tilde{a}(t)$ 分解, 得到 $\tilde{a}(t) = [\tilde{\varphi}(t), \tilde{\vartheta}(t)]$; 构造训练数据集 $\text{Angle_data} = \{\text{Angle_data}_i\}$, $\text{Angle_data}_i = \{a(t), \tilde{a}(t)\}$; 训练好的基于随机森林回归的飞行动作优化模型根据输入的动作 $a(t)$ 输出动作偏置 $\tilde{a}(t)$;

步骤4: 联合飞行状态优化模型、基于TD3的无人机航迹规划模型和基于随机森林回归的飞行动作优化模型, 对无人机航迹进行实时规划;

步骤4.1: 获取无人机携带的 I 个传感器在 t 时刻对无人机飞行状态的量测向量 $m_i(t)$, 输入至飞行状态优化模型中, 得到无人机 t 时刻飞行状态的最优估计 $\zeta(t|t) = [p(t), v(t)]^T$;

步骤4.2: 将无人机 t 时刻飞行状态的最优估计 $\zeta(t|t)$ 输入至训练好的基于TD3的无人机航迹规划模型中, 得到动作 $a(t) = [\varphi(t), \vartheta(t)]$;

步骤4.3: 将动作 $a(t)$ 输入至训练好的基于随机森林回归的飞行动作优化模型中, 得到动作偏置 $\tilde{a}(t) = [\tilde{\varphi}(t), \tilde{\vartheta}(t)]$;

步骤4.4: 计算 $\frac{a(t)}{\tilde{a}(t)} = \frac{\varphi(t)}{\tilde{\varphi}(t)} \times \frac{\vartheta(t)}{\tilde{\vartheta}(t)}$; 若 $\frac{a_i}{\tilde{a}_i} > 0.81$, 则直接输出动作 $a(t)$ 来控制无人机飞行;

否则, 执行步骤4.5;

步骤4.5: 优化飞行动作, 将原飞行动作 $a(t)$ 和飞行动作偏置 $\tilde{a}(t)$ 处理, 得到新的飞行动作 $a_{\text{new}}(t)$ 来控制无人机飞行:

$$a_{\text{new}}(t) = \begin{cases} [\varphi(t) + \tilde{\varphi}(t), \vartheta(t) + \tilde{\vartheta}(t)] & a(t) < \tilde{a}(t) \text{ and } a(t) + \pi/2 < \pi \\ [\max\{\varphi(t), \tilde{\varphi}(t)\} - \min\{\varphi(t), \tilde{\varphi}(t)\}, \max\{\vartheta(t), \tilde{\vartheta}(t)\} - \min\{\vartheta(t), \tilde{\vartheta}(t)\}] & a(t) > \tilde{a}(t) \text{ and } a(t) + \pi/2 < \pi \\ [\varphi(t) - \pi/2 - \tilde{\varphi}(t), \max\{\vartheta(t), \tilde{\vartheta}(t)\} - \min\{\vartheta(t), \tilde{\vartheta}(t)\}] & a(t) > \tilde{a}(t) \text{ and } a(t) + \pi/2 > \pi \\ [\varphi(t), \vartheta(t)] & \text{else} \end{cases}$$

步骤4.6: 在无人机的飞行过程中, 实时执行步骤4.1至步骤4.5, 实现每一时刻的前端飞行状态优化和后端飞行动作优化, 并实时控制优化后的飞行动作输出, 形成最终航迹。

一种基于联合优化的无人机航迹规划方法

技术领域

[0001] 本发明属于无人机飞行控制技术领域,具体涉及一种基于联合优化的无人机航迹规划方法。

背景技术

[0002] 无人机是一种不载人的通过无线传输遥控装置或自主传感控制设备完成相应任务的飞行设备,航迹规划是无人机执行飞行任务的有效技术手段,航迹可靠才能确保无人机完成飞行任务。

[0003] 传统的规划算法有人工势场法、Dijkstra算法等。人工势场法通过模拟引力场和排斥场对空间中无人机的综合作用规划无人机的飞行航迹,其适用于局部范围的规划,对全局规划的能力不足。Dijkstra算法是一种经典的用于最短路径求解的算法,它能够求出固定点到其他任意点的最短路径,简单有效,但缺点是一旦计算点数目增多,则算法计算量和所需内存极具增加。随着飞行环境日趋复杂,传统规划算法计算复杂度高、实时性差,难以实时控制无人机飞行。

[0004] 强化学习算法具有运算速度快、实时性强的特点,且能够根据无人机的飞行状态,端到端地决策飞行动作,使无人机飞行的实时控制成为可能,因此在规划领域广泛应用。通过对现有技术文献的检索发现,西北工业大学在其申请的专利“基于DDPG的无人机自主引导控制方法”(专利申请号:CN201910853746.X,申请公布号:CN110806756A)中提出了一种基于DDPG的无人机自主引导控制方法,该方法能够使无人机安全并快速地从起点飞到终点,提高了无人机执行任务的自主性和效率,但是该方法仅适用于静态地形环境。何金等在《兵工自动化》(2020,39(09):15-21)上发表的“未知环境下基于PF-DQN的无人机路径规划”中提出了一种环境信息未知情况下基于势函数奖赏的DQN路径规划方法,实现了无人机在环境信息未知下有效避障的路径规划,但是该算法的应用局限于与其训练近似的环境模型。已有文献的检索结果表明,这些方法的建模过程并未考虑实际环境中自然干扰因素的影响,如突发天气变化、风力、气流等,这类影响会导致无人机传感器采集的飞行数据具有偏差,使强化学习算法决策的飞行动作不佳,难以满足无人机飞行航迹的可靠性。

发明内容

[0005] 本发明的目的在于提供一种基于联合优化的无人机航迹规划方法。

[0006] 本发明的目的通过如下技术方案来实现:包括以下步骤:

[0007] 步骤1:搭建飞行状态优化模型;所述的飞行状态优化模型的输入为无人机携带的I个传感器在t时刻对无人机飞行状态的量测向量 $m_i(t)$,飞行状态优化模型的输出为无人机t时刻飞行状态的最优估计 $\zeta(t|t)=[p(t),v(t)]^T$;

[0008] 步骤1.1:输入无人机携带的I个传感器在t时刻对无人机飞行状态的量测向量 $m_i(t)$;

[0009] 其中,I个传感器的采样间隔均为 Ω ; $m_i(t)=[p_i(t),v_i(t)]^T$; $i=\{1,2,\dots,I\}$; p_i

(t) 表示第i个传感器获取的无人机位置量测向量, $p_i(t) = [p_{ix}(t), p_{iy}(t), p_{iz}(t)]^T$, $p_{ix}(t)$ 、 $p_{iy}(t)$ 和 $p_{iz}(t)$ 分别表示第i个传感器获取的笛卡尔坐标系下无人机在三维空间中x、y和z三个方向的位置坐标量测值; $v_i(t)$ 表示第i个传感器获取的无人机速度量测向量, $v_i(t) = [v_{ix}(t), v_{iy}(t), v_{iz}(t)]^T$, $v_{ix}(t)$ 、 $v_{iy}(t)$ 和 $v_{iz}(t)$ 分别表示第i个传感器获取的笛卡尔坐标系下无人机在三维空间中x、y和z三个方向的速度分量量测值;

[0010] 步骤1.2: 根据t-1时刻无人机飞行状态的最优估计 $\zeta(t-1|t-1)$, 计算t时刻无人机飞行状态的预测 $\zeta(t|t-1)$;

$$[0011] \quad \zeta(t|t-1) = \mathbf{F}\zeta(t-1|t-1) + \tilde{\omega}(t)$$

[0012] 其中, $\zeta(t-1|t-1) = [p(t-1), v(t-1)]^T$; $\mathbf{F} = \begin{bmatrix} 1 & \Omega \\ 1 & 1 \end{bmatrix}$; $\tilde{\omega}(t)$ 为白噪声;

[0013] 步骤1.3: 计算t时刻各传感器对无人机飞行状态的量测预测 $\delta_i(t|t-1)$;

$$[0014] \quad \delta_i(t|t-1) = \mathbf{H}_i(t)\zeta(t|t-1)$$

[0015] 其中, $\mathbf{H}_i(t)$ 表示第i个传感器的线性测量矩阵; $\mathbf{H}_i(t) = \begin{bmatrix} \frac{p_i(t)}{\|p_i(t)\|} & 1 \\ 1 & \frac{v_i(t)}{\|v_i(t)\|} \end{bmatrix}$; $\|\cdot\|$ 表示求模运算;

示求模运算;

[0016] 步骤1.4: 计算t时刻各传感器的量测新息 $\varepsilon_i(t)$;

$$[0017] \quad \varepsilon_i(t) = m_i(t) - \delta_i(t|t-1)$$

[0018] 步骤1.5: 根据t-1时刻无人机飞行状态的协方差 $P(t-1|t-1)$, 计算t时刻无人机飞行状态的协方差预测 $P(t|t-1)$;

$$[0019] \quad P(t|t-1) = \mathbf{F}P(t-1|t-1)\mathbf{F}^T + \mathbf{O}$$

$$[0020] \quad \text{其中, } \mathbf{O} = \begin{bmatrix} \frac{\Omega^3}{3} & \frac{\Omega^2}{2} \\ \frac{\Omega^2}{2} & \Omega \end{bmatrix};$$

[0021] 步骤1.6: 计算t时刻各传感器的预测新息协方差 $\mathbf{I}_i(t)$;

$$[0022] \quad \mathbf{I}_i(t) = \mathbf{H}_i(t)\mathbf{P}(t|t-1)\mathbf{H}_i^T(t) + \mathbf{R}_i(t)$$

$$[0023] \quad \text{其中, } \mathbf{R}_i(t) = \begin{bmatrix} 0.001 * m_i(t) & 0 \\ 0 & \frac{1}{m_i(t)} \end{bmatrix};$$

[0024] 步骤1.7: 计算t时刻各传感器量测向量的关联概率 $\beta_i(t)$;

$$[0025] \quad \beta_i(t) = \frac{e_i}{\sum_{i=1}^I e_i}$$

$$[0026] \quad e_i = \frac{1}{2\pi\|\mathbf{I}_i(t)\|_1^{\frac{1}{2}}} \exp\left[-\frac{1}{2}\varepsilon_i^T(t)\mathbf{I}_i^{-1}(t)\varepsilon_i(t)\right]$$

[0027] 步骤1.8:计算t时刻I个传感器的综合量测新息 $\varepsilon(t)$;

$$[0028] \quad \varepsilon(t) = \sum_{i=1}^I \beta_i(t) \varepsilon_i(t)$$

[0029] 步骤1.9:计算t时刻无人机飞行状态的协方差 $P(k|k)$;

$$[0030] \quad \mathbf{P}(t|t) = \mathbf{K}(t) \left[\sum_{i=1}^I \beta_i(t) \varepsilon_i(t) \varepsilon_i^T(t) - \varepsilon(t) \varepsilon^T(t) \right] \mathbf{K}^T(t)$$

$$[0031] \quad \mathbf{K}(t) = \sum_{i=1}^I \mathbf{P}(t|t-1) \mathbf{H}_i^T(t) \mathbf{I}_i^{-1}(t)$$

[0032] 步骤1.10:计算无人机t时刻飞行状态的最优估计 $\zeta(t|t)$;

$$[0033] \quad \zeta(t|t) = \zeta(t|t-1) \mathbf{K}(t) \varepsilon(t)$$

[0034] 步骤2:构建并训练基于TD3的无人机航迹规划模型;

[0035] 设置基于TD3的无人机航迹规划模型的状态空间输入为 $\zeta(t|t)$,设置基于TD3的无人机航迹规划模型的动作空间的输出为 $\mathbf{a}(t) = [\varphi(t), \vartheta(t)]$, $\varphi(t)$ 表示无人机t时刻飞行的方位角, $\vartheta(t)$ 表示无人机t时刻飞行的俯仰角;设置强化学习奖励函数 $r(t)$ 为:

$$[0036] \quad r(t) = r_1(t) + r_2(t)$$

$$[0037] \quad r_1(t) = \begin{cases} 1 - \|\mathbf{p}(t) - \mathbf{p}_{arrive}\| / 25 & \|\mathbf{p}(t) - \mathbf{p}_{arrive}\| < d_{max} \\ -\|\mathbf{p}(t) - \mathbf{p}_{arrive}\| / 25 & \|\mathbf{p}(t) - \mathbf{p}_{arrive}\| \geq d_{max} \end{cases}$$

$$[0038] \quad r_2(t) = \begin{cases} \|\mathbf{p}(t) - \mathbf{p}_{start}\| / 25 & \|\mathbf{p}(t) - \mathbf{p}_{start}\| < \rho_{max} \\ -2 & \|\mathbf{p}(t) - \mathbf{p}_{start}\| \geq \rho_{max} \end{cases}$$

[0039] 其中, r_1 表示到达正奖励; r_2 表示航程负奖励; \mathbf{p}_{arrive} 表示终点的位置坐标; \mathbf{p}_{start} 表示起点的位置坐标; d_{max} 表示无人机最大探测范围; ρ_{max} 表示无人机携带的燃料可供的最大可飞行航程;

[0040] 步骤3:构建并训练基于随机森林回归的飞行动作优化模型;

[0041] 利用步骤2中训练好的基于TD3的无人机航迹规划模型,根据每一时刻基于TD3的无人机航迹规划模型输出的动作 $\mathbf{a}(t) = [\varphi(t), \vartheta(t)]$,获取无人机执行动作 $\mathbf{a}(t)$ 后在t+1时刻到达的位置,将无人机t+1时刻的位置与t时刻的位置连接成标定线,标定线与无人机执行动作 $\mathbf{a}(t)$ 所得航迹构成夹角 $\tilde{\mathbf{a}}(t)$,在直角坐标系中 $\tilde{\mathbf{a}}(t)$ 分解,得到 $\tilde{\mathbf{a}}(t) = [\tilde{\varphi}(t), \tilde{\vartheta}(t)]$;构造训练数据集 $\text{Angle_data} = \{\text{Angle_data}_i\}$, $\text{Angle_data}_i = \{\mathbf{a}(t), \tilde{\mathbf{a}}(t)\}$;训练好的基于随机森林回归的飞行动作优化模型根据输入的动作 $\mathbf{a}(t)$ 输出动作偏置 $\tilde{\mathbf{a}}(t)$;

[0042] 步骤4:联合飞行状态优化模型、基于TD3的无人机航迹规划模型和基于随机森林回归的飞行动作优化模型,对无人机航迹进行实时规划;

[0043] 步骤4.1:获取无人机携带的I个传感器在t时刻对无人机飞行状态的量测向量 $\mathbf{m}_i(t)$,输入至飞行状态优化模型中,得到无人机t时刻飞行状态的最优估计 $\zeta(t|t) = [\mathbf{p}(t), \mathbf{v}(t)]^T$;

[0044] 步骤4.2:将无人机t时刻飞行状态的最优估计 $\zeta(t|t)$ 输入至训练好的基于TD3的

无人机航迹规划模型中,得到动作 $a(t)=[\varphi(t), \vartheta(t)]$;

[0045] 步骤4.3:将动作 $a(t)$ 输入至训练好的基于随机森林回归的飞行动作优化模型中,得到动作偏置 $\tilde{a}(t)=[\tilde{\varphi}(t), \tilde{\vartheta}(t)]$;

[0046] 步骤4.4:计算 $\frac{a(t)}{\tilde{a}(t)} = \frac{\varphi(t)}{\tilde{\varphi}(t)} \times \frac{\vartheta(t)}{\tilde{\vartheta}(t)}$;若 $\frac{a_t}{\tilde{a}_t} > 0.81$,则直接输出动作 $a(t)$ 来控制无人机飞行;否则,执行步骤4.5;

[0047] 步骤4.5:优化飞行动作,将原飞行动作 $a(t)$ 和飞行动作偏置 $\tilde{a}(t)$ 处理,得到新的飞行动作 a_{t_new} 来控制无人机飞行:

$$[0048] \quad a_{new}(t) = \begin{cases} [\varphi(t) + \tilde{\varphi}(t), \vartheta(t) + \tilde{\vartheta}(t)] & a(t) < \tilde{a}(t) \text{ and } a(t) + \pi/2 < \pi \\ [\max\{\varphi(t), \tilde{\varphi}(t)\} - \min\{\varphi(t), \tilde{\varphi}(t)\}, \max\{\vartheta(t), \tilde{\vartheta}(t)\} - \min\{\vartheta(t), \tilde{\vartheta}(t)\}] & a(t) > \tilde{a}(t) \text{ and } a(t) + \pi/2 < \pi \\ [\varphi(t) - \pi/2 - \tilde{\varphi}(t), \max\{\vartheta(t), \tilde{\vartheta}(t)\} - \min\{\vartheta(t), \tilde{\vartheta}(t)\}] & a(t) > \tilde{a}(t) \text{ and } a(t) + \pi/2 > \pi \\ [\varphi(t), \vartheta(t)] & \text{else} \end{cases}$$

[0049] 步骤4.6:在无人机的飞行过程中,实时执行步骤4.1至步骤4.5,实现每一时刻的前端飞行状态优化和后端飞行动作优化,并实时控制优化后的飞行动作输出,形成最终航迹。

[0050] 本发明的有益效果在于:

[0051] 本发明通过设计前端飞行状态优化模型,引入目标跟踪机理,最优估计无人机自身的飞行状态,大幅度降低了自然干扰因素对传感器数据的误差;通过设计后端飞行动作优化模型,结合姿态优化机制,将实际偏差大的飞行角度调整至预期飞行角度,解决了强化学习决策飞行动作不佳的问题,提高了飞行航迹的可靠性。

附图说明

[0052] 图1是本发明的总体流程图。

[0053] 图2是本发明的实施例中基于PDA的无人机飞行状态优化方法的流程图。

[0054] 图3是本发明的实施例中基于TD3的无人机航迹规划方法训练流程图。

[0055] 图4是无人机球面坐标系下速度模型图。

[0056] 图5是本发明的实施例中动作优化回归模型数据集构建方法的示意图。

[0057] 图6是本发明的实施例中应用基于联合优化的无人机航迹规划方法的流程图。

具体实施方式

[0058] 下面结合附图对本发明做进一步描述。

[0059] 本发明属于无人机飞行控制领域,涉及一种基于联合优化的无人机航迹规划方法。本发明降低了自然干扰因素对无人机飞行航迹的影响,联合前端飞行状态优化和后端飞行动作优化。本发明的技术方案具体包括以下步骤:

[0060] 步骤1:搭建飞行状态优化模型,利用概率数据关联算法实现无人机传感器量测数据到最优飞行状态的估计;

[0061] 步骤2:结合步骤1的最优飞行状态,构建基于TD3的无人机航迹规划模型,实现无人机对规划环境的探索;

[0062] 步骤3:结合步骤2的TD3模型,设计基于随机森林回归的飞行动作优化模型,实现飞行动作到飞行动作修正偏置的映射;

[0063] 步骤4:联合步骤1的飞行状态优化模型和步骤2、步骤3的飞行动作优化模型,构成基于联合优化的无人机航迹规划模型,实时控制飞行动作输出,形成航迹,

[0064] 步骤1包括:

[0065] 选择概率数据关联(Probability Data Association,PDA)算法为状态优化算法。

[0066] 将无人机在实际环境中飞行受到的自然干扰因素扰动作为控制因素 $\tilde{\omega}(k)$,结合无人机飞行的机动模型,获得飞行状态模型预测。

[0067] 根据实际情况设定传感器线性量测矩阵 $\mathbf{H}(k) = \begin{bmatrix} \frac{p(k)}{\|p(k)\|} & 1 \\ 1 & \frac{v(k)}{\|v(k)\|} \end{bmatrix}$,以获得飞行状态

的量测预测。式中, $p(k) = [p_x(k), p_y(k), p_z(k)]^T$ 表示无人机位置信息, $p_x(k)$ 、 $p_y(k)$ 和 $p_z(k)$ 分别代表笛卡尔坐标系下,无人机在三维空间中x、y和z三个方向的位置坐标点, $v(k) = [v_x(k), v_y(k), v_z(k)]^T$ 表示k时刻无人机速度信息, $v_x(k)$ 、 $v_y(k)$ 和 $v_z(k)$ 分别代表笛卡尔坐标系下,无人机在三维空间中x、y和z三个方向的速度分量, $\|\cdot\|$ 表示求模运算。

[0068] 利用机载传感器,量测得到多组无人机自身飞行状态数据,将所得数据均作为有效量测数据,结合量测预测,获得每个传感器量测数据的新息。

[0069] 将每个传感器量测数据的关联概率和新息加权计算,得到综合新息,并结合卡尔曼滤波器增益和飞行状态模型预测,得到最优估计的飞行状态。

[0070] 步骤2包括:

[0071] 选择双延迟深度确定性策略梯度(Twin Delayed Deep Deterministic Policy Gradient,TD3)算法为无人机航迹规划算法。

[0072] 考虑由PDA得到的最优飞行状态,设计无人机航迹规划的强化学习状态空间为 $s_t = [p_t, v_t]^T$,式中, $p_t = [p_{x,t}, p_{y,t}, p_{z,t}]^T$ 表示t时刻无人机位置信息, $p_{x,t}$ 、 $p_{y,t}$ 和 $p_{z,t}$ 分别代表笛卡尔坐标系下,无人机在三维空间中x、y和z三个方向的位置坐标点, $v_t = [v_{x,t}, v_{y,t}, v_{z,t}]^T$ 为t时刻无人机速度信息, $v_{x,t}$ 、 $v_{y,t}$ 和 $v_{z,t}$ 分别代表笛卡尔坐标系下,无人机在三维空间中x、y和z三个方向的速度分量。

[0073] 考虑无人机匀速率飞行,控制其速度方向就可以控制飞行的方向,设计无人机航迹规划的强化学习动作空间为 $a_t = [\varphi_t, \vartheta_t]$,式中, φ_t 和 ϑ_t 分别表示无人机飞行的方位角和俯仰角。

[0074] 考虑模型训练应使无人机飞行趋向于终点并避免燃料耗尽,设计无人机航迹规划的强化学习奖励函数为 $r_t = r_1 + r_2$,式中, r_1 表示到达正奖励, r_2 表示航程负奖励,

$$r_1 = \begin{cases} 1 - \|p_t - p_{arrive}\| / 25, & \text{if } \|p_t - p_{arrive}\| < d_{max} \\ -\|p_t - p_{arrive}\| / 25, & \text{else} \end{cases}, \text{ 式中, } p_t \text{ 表示 } t \text{ 时刻无人机所处位置坐标, } p_{arrive}$$

表示终点位置坐标, d_{max} 表示无人机最大探测范围; $r_2 = \begin{cases} \|p_t - p_{start}\| / 25, & \text{if } \|p_t - p_{start}\| < \rho_{max} \\ -2, & \text{else} \end{cases}$,

式中, p_{start} 表示起点位置坐标, p_{max} 表示无人机最大可飞行航程。

[0075] 搭建TD3网络结构,其中Actor网络和Critic网络,均为3层全连接网络结构,Actor网络的输入层为2个神经元,输出层为6个神经元,Critic网络的输入层为8个神经元,输出层为1个神经元,设置TD3训练参数,进行网络训练,得到基于TD3的无人机航迹规划模型。

[0076] 步骤3包括:

[0077] 定义回归模型训练的数据集 $Angle_data = \{Angle_data_t\}$, 式中, $Angle_data_t = \{a_t, \tilde{a}_t\}$ 为任意t时刻的角度数据集, $a_t = [\varphi_t, \vartheta_t]$ 表示无人机的飞行动作, 作为样本, $\tilde{a}_t = [\tilde{\varphi}_t, \tilde{\vartheta}_t]$ 表示无人机航迹连线和无人机与终点连线的夹角, 即飞行动作偏置, 作为标签, $\tilde{\varphi}_t$ 和 $\tilde{\vartheta}_t$ 为 \tilde{a}_t 在直角坐标系的方位角和俯仰角。

[0078] 利用步骤2得到的已训练TD3模型的Actor在线网络, 决策无人机在规划空间中每一时刻的飞行动作 $a_t = [\varphi_t, \vartheta_t]$; 无人机执行飞行动作, 飞行到下一位置; 将该步飞行前无人机所处位置和无人机任务终点形成标定线, 标定线与该步飞行所得航迹构成夹角 \tilde{a}_t ; 在直角坐标系中对该夹角分解, 得到 $\tilde{a}_t = [\tilde{\varphi}_t, \tilde{\vartheta}_t]$; 得到每一时刻的 $Angle_data_t$ 并存储; 完成整个飞行过程, 存储整体数据集 $Angle_data$ 。

[0079] 设置随机森林回归模型的基本参数, 结合数据集 $Angle_data$, 进行回归模型训练, 得到基于随机森林回归的飞行动作优化模型。

[0080] 步骤4包括:

[0081] 无人机携带4个传感器实时量测飞行状态数据, 利用PDA算法得到无人机在该时刻的最优飞行状态。

[0082] 最优飞行状态作为步骤2所得已训练TD3模型的输入, 利用其中的Actor在线网络, 决策飞行动作输出。

[0083] 将飞行动作, 作为步骤3所得已训练随机森林回归模型的输入, 得到对应的飞行动作偏置。

[0084] 判决飞行动作偏置, 定义 $\frac{a_t}{\tilde{a}_t} = \frac{\varphi_t}{\tilde{\varphi}_t} \times \frac{\vartheta_t}{\tilde{\vartheta}_t}$, 若 $\frac{a_t}{\tilde{a}_t} > 0.81$, 则说明飞行航迹与标定航迹(标

定航迹即无人机所在位置和终点的连线) 接近, 不进行动作优化; 若 $\frac{a_t}{\tilde{a}_t} \leq 0.81$, 则说明飞行航迹偏离标定航迹, 需要进行动作优化。

[0085] 依照判决结果, 若不优化飞行动作, 则将已训练TD3模型所得飞行动作控制输出; 若需要优化飞行动作, 则将原飞行动作 a_t 和飞行动作偏置 \tilde{a}_t 依照式(1)处理, 得到新的飞行动作 a_{t_new} , 并控制飞行动作实时输出, 形成最终航迹。

$$[0086] \quad a_{t_new} = \begin{cases} a_t + \tilde{a}_t & , \text{if } a_t < \tilde{a}_t \text{ and } a_t + \pi/2 < \pi \\ a_t - \tilde{a}_t & , \text{if } a_t > \tilde{a}_t \text{ and } a_t + \pi/2 < \pi \\ a_t - \pi/2 - \tilde{a}_t & , \text{if } a_t > \tilde{a}_t \text{ and } a_t + \pi/2 > \pi \\ a_t & , \text{else} \end{cases} \quad (1)$$

[0087] 本发明的有益效果在于:

[0088] 本发明设计的前端飞行状态优化模型,引入了目标跟踪机理,最优估计无人机自身的飞行状态,大幅度降低了自然干扰因素对传感器数据的误差;后端飞行动作优化模型结合了姿态优化机制,将实际偏差大的飞行角度调整至预期飞行角度,解决了强化学习决策飞行动作不佳的问题,提高了飞行航迹的可靠性。总之,本发明联合前端飞行状态优化和后端飞行动作优化,提供的基于联合优化的无人机航迹规划方法,在自然干扰因素的影响下,提高了无人机飞行航迹的可靠性,提升了无人机抗干扰的能力。

[0089] 实施例1:

[0090] 本发:针对实际环境中,无人机传感器采集的数据受到自然干扰影响存在误差,导致强化学习算法决策飞行动作不佳的问题,提供一种基于联合优化的无人机航迹规划方法。本发明通过搭建基于概率数据关联的飞行状态优化模型,完成无人机传感器量测数据到飞行状态的最优估计;设定无人机航迹规划背景下的强化学习基本要素,结合最优飞行状态设计状态空间,依据无人机机动原理设计动作空间,根据实际任务需求设计奖励函数,构建并训练基于TD3的无人机航迹规划模型,实时决策飞行动作;利用已训练TD3,构建动作优化回归模型数据集,设计并训练基于随机森林回归的飞行动作优化模型,完成飞行动作到飞行动作修正偏置的映射。在实际应用中,对于传感器实时采集飞行状态数据,利用状态优化模型实时获得最优飞行状态,将其作为已训练TD3模型的输入,得到飞行动作,并利用飞行动作优化模型,得到对应的飞行动作偏置,对其进行判决,控制输出优化后的飞行动作,从而在联合优化的作用下,实现无人机面对自然干扰影响的实时航迹规划。

[0091] 如附图1所示是本发明实施例基于联合优化的无人机航迹规划方法流程示意图,包括步骤1至步骤4:

[0092] 步骤1:搭建飞行状态优化模型,利用概率数据关联算法实现无人机传感器量测数据到最优飞行状态的估计。

[0093] 如附图2所示是本发明实施例基于PDA的无人机飞行状态优化方法流程示意图,包括步骤1.1至步骤1.13:

[0094] 步骤1.1:设计在任意k时刻,无人机运动的状态向量为:

$$[0095] \quad \zeta(k) = [p(k), v(k)]^T \quad (2)$$

[0096] 式中, $p(k) = [p_x(k), p_y(k), p_z(k)]^T$ 表示无人机位置信息, $p_x(k)$ 、 $p_y(k)$ 和 $p_z(k)$ 分别代表笛卡尔坐标系下,无人机在三维空间中x、y和z三个方向的位置坐标点, $v(k) = [v_x(k), v_y(k), v_z(k)]^T$ 表示k时刻无人机速度信息, $v_x(k)$ 、 $v_y(k)$ 和 $v_z(k)$ 分别代表笛卡尔坐标系下,无人机在三维空间中x、y和z三个方向的速度分量。

[0097] 步骤1.2:根据k-1时刻无人机的运动状态 $\zeta(k-1|k-1)$,完成k时刻的状态预测 $\zeta(k|k-1)$:

$$[0098] \quad \zeta(k|k-1) = \mathbf{F}(k-1)\zeta(k-1|k-1) + \tilde{\omega}(k) \quad (3)$$

[0099] 式中, $\mathbf{F} = \begin{bmatrix} 1 & \Omega \\ 1 & 1 \end{bmatrix}$ 为任意 Ω 个采样间隔内无人机的机动模型,在本发明的实施例中,

无人机飞行采用匀速运动模型, $\tilde{\omega}(k)$ 为无人机在实际环境中飞行的随机扰动,在本发明的实施例中,将其设计为白噪声形式。

[0100] 步骤1.3:计算k时刻的量测预测 $\delta(k|k-1)$:

[0101] $\delta(k|k-1) = H(k) \zeta(k|k-1)$ (4)

[0102] 式中, $H(k)$ 表示线性测量矩阵, 在本发明的实施例中, 将其设定为:

[0103]
$$\mathbf{H}(k) = \begin{bmatrix} \frac{p(k)}{\|p(k)\|} & 1 \\ 1 & \frac{v(k)}{\|v(k)\|} \end{bmatrix} \quad (5)$$

[0104] 式中, $\|\cdot\|$ 表示求模运算。

[0105] 步骤1.4: 利用传感器, 获得有效量测数据 $M(k)$, 在本发明的实施例中, 无人机携带4个传感器:

[0106] $M(k) = m_i(k), i=1, 2, 3, 4$ (6)

[0107] 式中, $m_i(k)$ 即为在 k 时刻, 第 i 个传感器得到的状态的有效量测值, 在本发明的实施例中, 规定传感器量测的数据均为有效量测值。

[0108] 步骤1.5: 计算每个有效量测值 $m_i(k)$ 的新息:

[0109] $\varepsilon_i(k) = m_i(k) - \delta(k|k-1), i=1, 2, 3, 4$ (7)

[0110] 式中, $\varepsilon_i(k)$ 即为在 k 时刻, 第 i 个传感器所得量测值的新息。

[0111] 步骤1.6: 计算 $k-1$ 时刻无人机状态协方差为:

[0112]
$$\mathbf{P}(k-1) = \begin{bmatrix} \text{Cov}(p(k-1), p(k-1)) & \text{Cov}(p(k-1), v(k-1)) \\ \text{Cov}(v(k-1), p(k-1)) & \text{Cov}(v(k-1), v(k-1)) \end{bmatrix} \quad (8)$$

[0113] 式中, $\text{Cov}(\cdot, \cdot)$ 表示协方差运算。

[0114] 步骤1.7: 计算 k 时刻的状态协方差预测 $P(k|k-1)$:

[0115] $P(k|k-1) = F(k-1)P(k-1|k-1)F^T(k-1) + \mathbf{O}$ (9)

[0116] 式中, $\mathbf{O} = \begin{bmatrix} \frac{\Omega^3}{3} & \frac{\Omega^2}{2} \\ \frac{\Omega^2}{2} & \Omega \end{bmatrix}$ 为本发明的实施例中, 白噪声形式的随机扰动对应的协方差。

[0117] 步骤1.8: 预测新息的协方差 $I(k)$:

[0118] $I(k) = H(k)P(k|k-1)H^T(k) + R(k)$ (10)

[0119] 式中, $R(k)$ 为量测噪声协方差, 在本发明的实施例中, 将其设定为:

[0120]
$$\mathbf{R}(k) = \begin{bmatrix} 0.001 * \mathbf{M}^2(k) & \mathbf{0} \\ \mathbf{0} & \frac{1}{\mathbf{M}(k)} \end{bmatrix} \quad (11)$$

[0121] 步骤1.9: 计算每个量测值的关联概率 $\beta_i(k)$:

[0122]
$$\beta_i(k) = \frac{e_i}{\sum_{i=1}^4 e_i}, i=1, 2, 3, 4 \quad (12)$$

[0123] 式中, $e_i = \frac{1}{2\pi \|\mathbf{I}(k)\|_1^{\frac{1}{2}}} \exp\left[-\frac{1}{2} \boldsymbol{\varepsilon}_i^T(k) \mathbf{I}^{-1}(k) \boldsymbol{\varepsilon}_i(k)\right], i=1, 2, 3, 4。$

[0124] 步骤1.10:依据公式(13),综合新息 $\varepsilon_i(k)$:

$$[0125] \quad \varepsilon(k) = \sum_{i=1}^4 \beta_i(k) \varepsilon_i(k) \quad (13)$$

[0126] 步骤1.11:更新无人机在k时刻的最优状态 $\zeta(k|k)$,用于k+1时刻的最优飞行状态更新:

$$[0127] \quad \zeta(k|k) = \zeta(k|k-1) + K(k) \varepsilon(k) \quad (14)$$

[0128] 式中, $K(k) = P(k|k-1) H^T(k) I^{-1}(k)$ 为卡尔曼滤波器增益。

[0129] 步骤1.12:更新无人机在k时刻的状态协方差预测 $P(k|k)$,用于k+1时刻的最优飞行状态更新:

$$[0130] \quad P(k|k) = K(k) \left[\sum_{i=1}^4 \beta_i(k) \varepsilon_i(k) \varepsilon_i^T(k) - \varepsilon(k) \varepsilon^T(k) \right] K^T(k) \quad (15)$$

[0131] 步骤1.13:在无人机的飞行的每一时刻均执行步骤1.2至步骤1.12,实现无人机飞行状态在每一时刻由传感器量测数据到最优飞行状态的估计。

[0132] 步骤2:结合步骤1的最优飞行状态,构建基于TD3的无人机航迹规划模型,实现无人机对规划环境的探索。

[0133] 如附图3所示是本发明实施例基于TD3的无人机航迹规划方法训练流程示意图,包括步骤2.1至步骤2.6:

[0134] 步骤2.1:在本发明的实施例中,选择TD3作为实现无人机航迹规划的算法。

[0135] 步骤2.2:在本发明的实施例中,考虑由PDA得到的最优飞行状态 $\zeta(k|k) = [p(k), v(k)]^T$,设计无人机航迹规划的强化学习状态空间 s_t 为:

$$[0136] \quad s_t = [p_t, v_t]^T = \zeta(t|t) \quad (16)$$

[0137] 式中, $p_t = [p_{x,t}, p_{y,t}, p_{z,t}]^T$ 表示t时刻无人机位置信息, $p_{x,t}$ 、 $p_{y,t}$ 和 $p_{z,t}$ 分别代表笛卡尔坐标系下,无人机在三维空间中x、y和z三个方向的位置坐标点, $v_t = [v_{x,t}, v_{y,t}, v_{z,t}]^T$ 为t时刻无人机速度信息, $v_{x,t}$ 、 $v_{y,t}$ 和 $v_{z,t}$ 分别代表笛卡尔坐标系下,无人机在三维空间中x、y和z三个方向的速度分量。

[0138] 图4所示是无人机球面坐标系下速度模型图,如附图4所示,1是无人机在空间中飞行的点迹,2是坐标系的x轴方向,3是坐标系的y轴方向,4是坐标系的z轴方向,5是无人机飞行速度方向,6是无人机飞行速度方向的方位角 φ_t ,7是无人机飞行速度方向的俯仰角 θ_t 。

[0139] 步骤2.3:在本发明的实施例中,设计无人机航迹规划的强化学习动作空间 a_t 为:

$$[0140] \quad a_t = [\varphi_t, \theta_t] \quad (17)$$

[0141] 式中, φ_t 和 θ_t 分别表示无人机飞行的方位角和俯仰角。

[0142] 步骤2.4:在本发明的实施例中,设计无人机航迹规划的强化学习奖励函数 r_t 为:

$$[0143] \quad r_t = r_1 + r_2 \quad (18)$$

[0144] 式中, r_1 表示到达正奖励, r_2 表示航程负奖励。

[0145] r_1 具体设计为:

$$[0146] \quad r_1 = \begin{cases} 1 - \|p_t - p_{arrive}\| / 25, & \text{if } \|p_t - p_{arrive}\| < d_{max} \\ -\|p_t - p_{arrive}\| / 25, & \text{else} \end{cases} \quad (19)$$

[0147] 式中, p_t 表示t时刻无人机所处位置坐标, p_{arrive} 表示终点位置坐标, d_{max} 表示无人机最大探测范围。

[0148] r_2 具体设计为:

$$[0149] \quad r_2 = \begin{cases} \|p_t - p_{start}\| / 25, & \text{if } \|p_t - p_{start}\| < \rho_{max} \\ -2, & \text{else} \end{cases} \quad (20)$$

[0150] 式中, p_{start} 表示起点位置坐标, ρ_{max} 表示无人机携带的燃料可供的最大可飞行航程。

[0151] 步骤2.5: 搭建Actor网络和Critic网络, 均为3层全连接网络结构, Actor网络的输入层为2个神经元, 输出层为6个神经元, Critic网络的输入层为8个神经元, 输出层为1个神经元。设置训练参数: 回合更新次数为10000, 每回合时间步数为50, 延迟步数为2, 记忆池大小为5000, 采集经验数据的批量大小为32, 折扣因子大小为0.99, Actor网络学习率为 1×10^{-4} , Critic网络学习率为 2×10^{-4} 。

[0152] 步骤2.6: 利用步骤2.1至2.4设置的无人机航迹规划背景下的强化学习状态空间、动作空间和奖励函数, 在步骤2.5设定的参数条件下, 依据TD3的原理, 训练基于TD3的无人机航迹规划模型, 获得无人机从飞行状态到飞行动作端到端的决策映射。

[0153] 步骤3: 结合步骤2的TD3模型, 设计基于随机森林回归的飞行动作优化模型, 实现飞行动作到飞行动作修正偏置的映射。

[0154] 图5是本发明实施例的动作优化回归模型数据集构建方法示意图, 如附图5所示, 8是无人机飞行空间直角坐标系原点, 9是直角坐标系x轴方向, 10是直角坐标系y轴方向, 11是直角坐标系z轴方向, 12是无人机飞行的起点即航迹的第0点, 13是1时刻的航迹点, 14是2时刻的航迹点, 15是t时刻的航迹点, 16是t+1时刻的航迹点, 17是无人机飞行的终点, 18是空间中的障碍, 19是无人机飞行起点与1时刻航迹点形成的航迹连线和起点至终点的标定线的夹角 \tilde{a}_0 , 20是无人机1时刻航迹点与2时刻航迹点形成的航迹连线和1时刻航迹点至终点的标定线的夹角 \tilde{a}_1 , 21是无人机t时刻航迹点与t+1时刻航迹点形成的航迹连线和t时刻航迹点至终点的标定线的夹角 \tilde{a}_t 。

[0155] 步骤3.1: 在本发明的实施例中, 生成回归模型训练的数据集Angle_data:

$$[0156] \quad \text{Angle_data} = \{\text{Angle_data}_t\} \quad (21)$$

$$[0157] \quad \text{Angle_data}_t = \{a_t, \tilde{a}_t\} \quad (22)$$

[0158] 式中, $\text{Angle_data}_t = \{a_t, \tilde{a}_t\}$ 为任意t时刻的角度数据集, $a_t = [\varphi_t, \mathcal{G}_t]$ 表示无人机的飞行动作, 作为样本, $\tilde{a}_t = [\tilde{\varphi}_t, \tilde{\mathcal{G}}_t]$ 表示无人机航迹连线和无人机与终点连线的夹角, 即飞行动作偏置, 作为标签, $\tilde{\varphi}_t$ 和 $\tilde{\mathcal{G}}_t$ 为 \tilde{a}_t 在直角坐标系的方位角和俯仰角。

[0159] 步骤3.2: 在本发明的实施例中, 利用步骤2已训练TD3模型的Actor在线网络, 在图5所示规划空间中决策无人机每一时刻的飞行动作, 即 $a_t = [\varphi_t, \mathcal{G}_t]$; 无人机执行飞行动作, 飞行到下一位置; 将该步飞行前无人机所处位置和终点形成标定线, 标定线与该步飞行所得航迹构成夹角 \tilde{a}_t ; 在直角坐标系中对该夹角分解, 得到 $\tilde{a}_t = [\tilde{\varphi}_t, \tilde{\mathcal{G}}_t]$; 得到每一时刻的Angle_data_t 并存储; 完成整个飞行过程, 存储整体数据集Angle_data。

[0160] 步骤3.3:在本发明的实施例中,设置随机森林回归模型的基本参数:决策树个数100,决策树最大深度10。

[0161] 步骤3.4:利用步骤3.2得到的关于步骤3.1生成的数据集Angle_data,在步骤3.3设定的参数条件下,依据随机森林回归的原理,训练随机森林回归模型,得到飞行动作到飞行动作修正偏置的映射。

[0162] 步骤4:联合步骤1的飞行状态优化模型和步骤2、步骤3的飞行动作优化模型,构成基于联合优化的无人机航迹规划模型,实时控制飞行动作输出,形成航迹。

[0163] 如附图6所示是本发明实施例应用基于联合优化的无人机航迹规划方法示意图,包括步骤4.1至步骤4.6:

[0164] 步骤4.1:对于4个传感器实时量测的4组状态数据,利用PDA算法实施步骤1,得到无人机在该时刻的最优飞行状态 $\zeta(t|t)$ 。

[0165] 步骤4.2:将步骤4.1获得的最优飞行状态 $\zeta(t|t)$,作为步骤2所得已训练TD3模型的输入 s_t ,利用其中的Actor在线网络,决策飞行动作 a_t 。

[0166] 步骤4.3:将步骤4.2获得的飞行动作 a_t ,作为步骤3所得已训练随机森林回归模型的输入,得到对应的飞行动作偏置 \tilde{a}_t 。

[0167] 步骤4.4:在本发明的实施例中,定义判决式(23),并判决飞行动作偏置:

$$[0168] \quad \frac{a_t}{\tilde{a}_t} = \frac{\varphi_t}{\tilde{\varphi}_t} \times \frac{\mathcal{G}_t}{\tilde{\mathcal{G}}_t} \quad (23)$$

[0169] 若 $\frac{a_t}{\tilde{a}_t} > 0.81$,则说明飞行航迹与标定航迹(标定航迹即无人机所在位置和终点的连线)接近,不进行动作优化;若 $\frac{a_t}{\tilde{a}_t} \leq 0.81$,则说明飞行航迹偏离标定航迹,需要进行动作优化。

[0170] 步骤4.5:依照步骤4.4的判决结果,若不优化飞行动作,则将步骤4.2所得飞行动作 a_t 控制输出;若需要优化飞行动作,则将原飞行动作 a_t 和飞行动作偏置 \tilde{a}_t 依照式(24)及式(25)处理,得到新的飞行动作 a_{t_new} :

$$[0171] \quad a_{t_new} = \begin{cases} a_t + \tilde{a}_t, & \text{if } a_t < \tilde{a}_t \text{ and } a_t + \pi/2 < \pi \\ a_t - \tilde{a}_t, & \text{if } a_t > \tilde{a}_t \text{ and } a_t + \pi/2 < \pi \\ a_t - \pi/2 - \tilde{a}_t, & \text{if } a_t > \tilde{a}_t \text{ and } a_t + \pi/2 > \pi \\ a_t, & \text{else} \end{cases} \quad (24)$$

$$[0172] \quad \begin{cases} a_t + \tilde{a}_t = & [\varphi_t + \tilde{\varphi}_t, \mathcal{G}_t + \tilde{\mathcal{G}}_t] \\ a_t - \tilde{a}_t = & [\max\{\varphi_t, \tilde{\varphi}_t\} - \min\{\varphi_t, \tilde{\varphi}_t\}, \max\{\mathcal{G}_t, \tilde{\mathcal{G}}_t\} - \min\{\mathcal{G}_t, \tilde{\mathcal{G}}_t\}] \\ a_t - \pi/2 - \tilde{a}_t = & [\varphi_t - \pi/2 - \tilde{\varphi}_t, \max\{\mathcal{G}_t, \tilde{\mathcal{G}}_t\} - \min\{\mathcal{G}_t, \tilde{\mathcal{G}}_t\}] \\ a_t = & [\varphi_t, \mathcal{G}_t] \end{cases} \quad (25)$$

[0173] 步骤4.6:在无人机的飞行过程中,实时执行步骤4.1至步骤4.5,实现每一时刻的前端飞行状态优化和后端飞行动作优化,并实时控制优化后的飞行动作输出,形成最终航迹。

[0174] 以上所述仅为本发明的优选实施例而已,并不用于限制本发明,对于本领域的技

术人员来说,本发明可以有各种更改和变化。凡在本发明的精神和原则之内,所作的任何修改、等同替换、改进等,均应包含在本发明的保护范围之内。

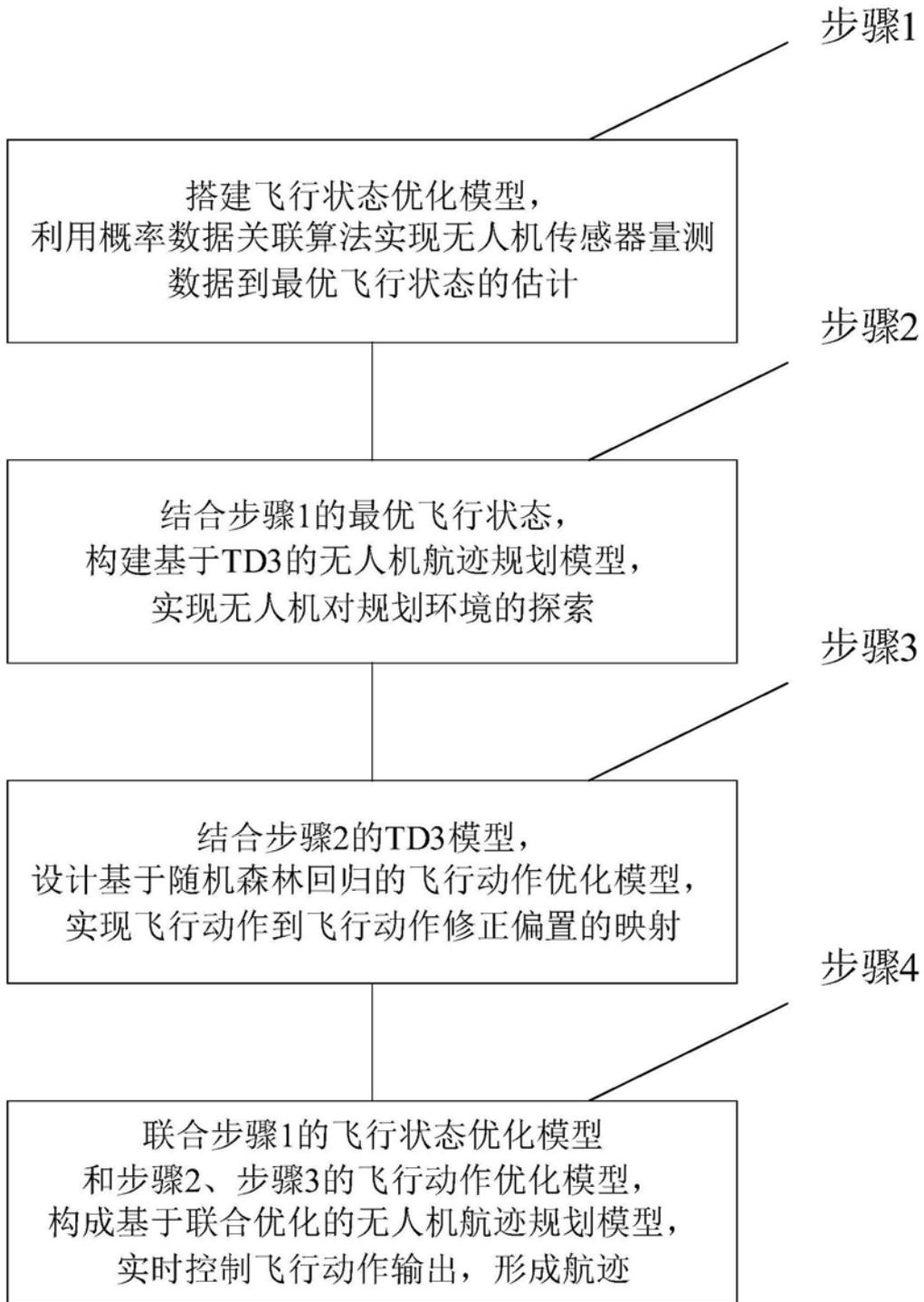


图1

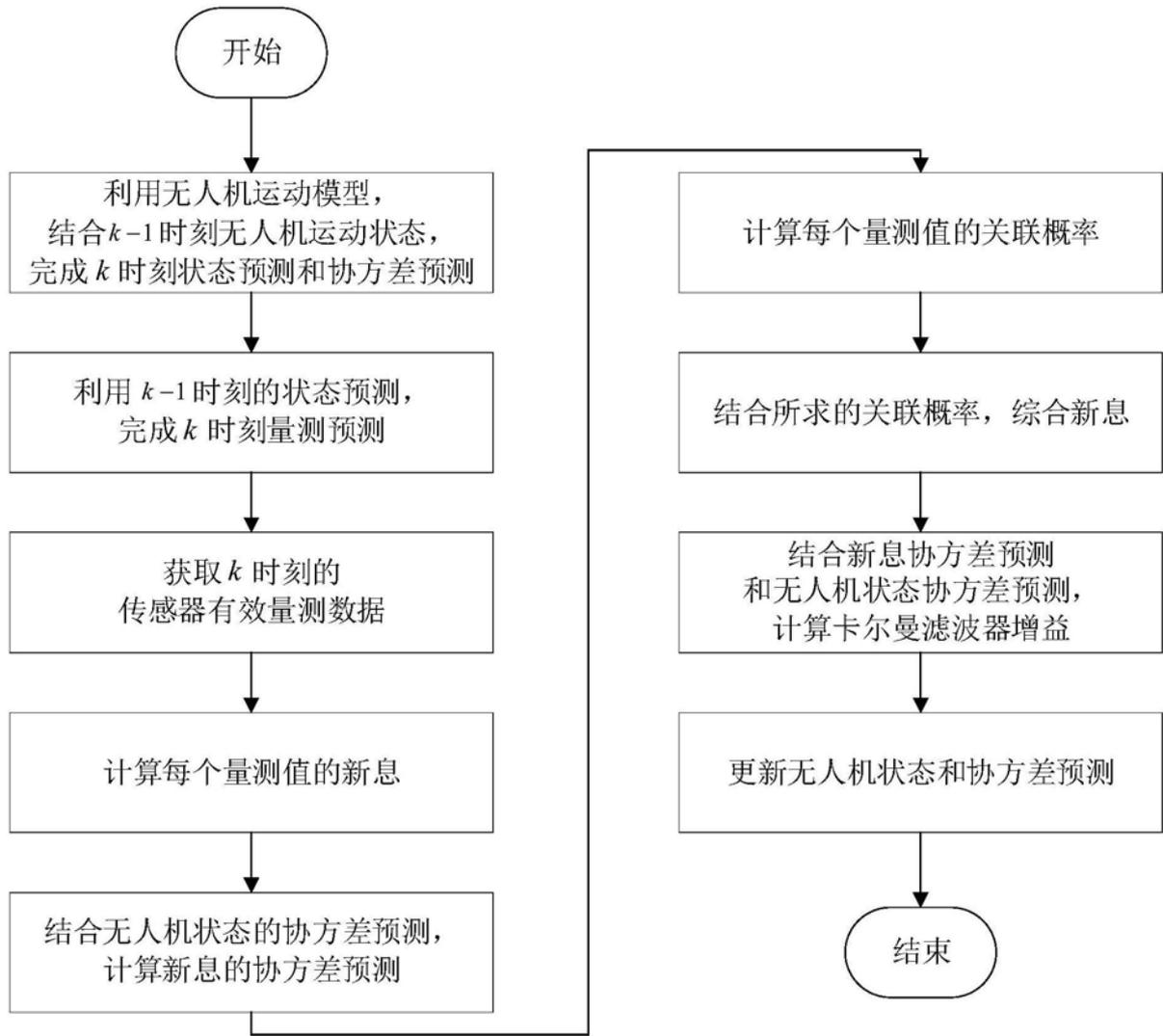


图2

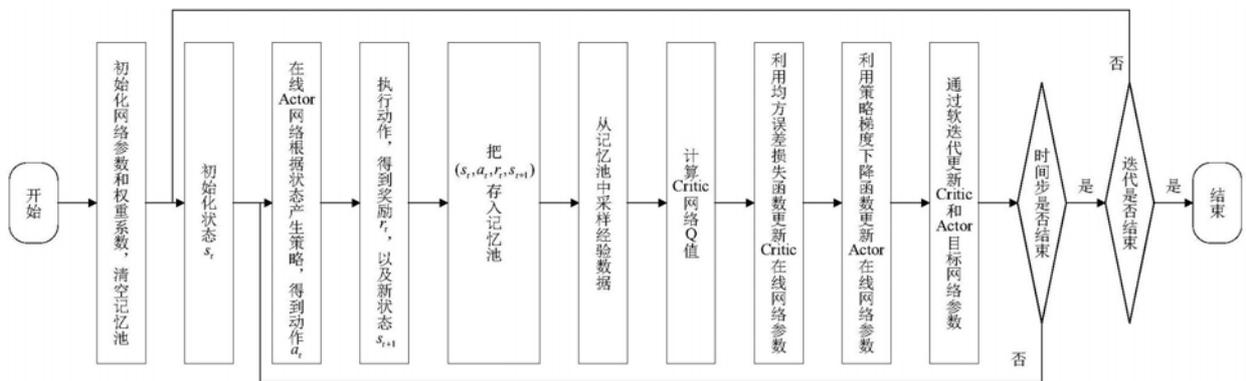


图3

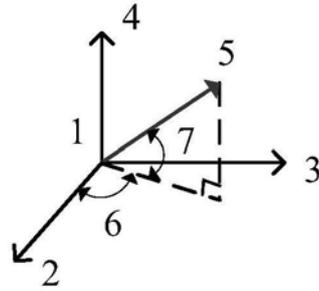


图4

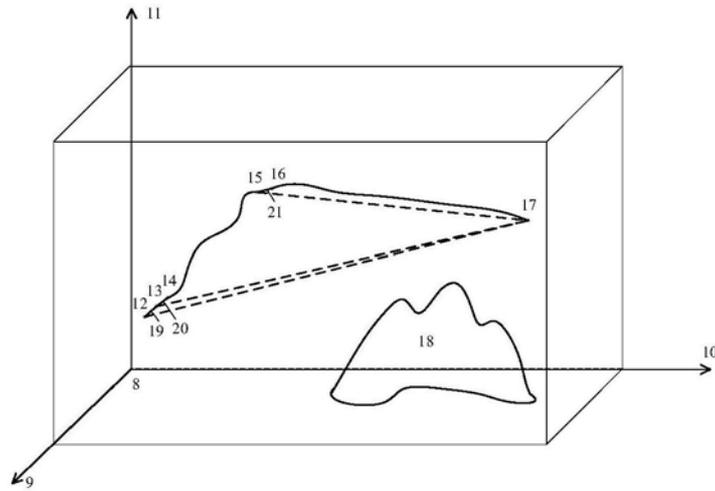


图5

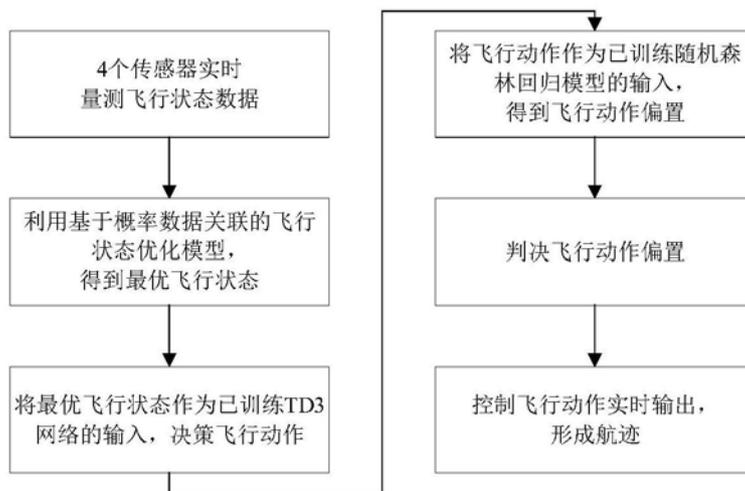


图6