



基于概率推理学习优化的无人自行车质量偏心校正方法

黄用华, 梁子彦, 庄未, 任仰华

引用本文:

黄用华, 梁子彦, 庄未, 等. 基于概率推理学习优化的无人自行车质量偏心校正方法[J]. 控制与决策, 2025, 40(5): 1723–1732.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2024.0775>

您可能感兴趣的其他文章

Articles you may be interested in

基于数据分布特性的代价敏感宽度学习系统

[Data distribution-based cost-sensitive broad learning system](#)

控制与决策. 2021, 36(7): 1686–1692 <https://doi.org/10.13195/j.kzyjc.2019.1484>

基于MCPDDPG的智能车辆路径规划方法及应用

The method and application of intelligent vehicle path planning based on MCPDDPG

控制与决策. 2021, 36(4): 835–846 <https://doi.org/10.13195/j.kzyjc.2019.0460>

基于FWADE-ELM的短时交通流预测方法

Short-term traffic flow forecasting based on hybrid FWADE-ELM

控制与决策. 2021, 36(4): 925–932 <https://doi.org/10.13195/j.kzyjc.2019.1103>

基于Frenet坐标系的自动驾驶轨迹规划与优化算法

Trajectory planning and optimization algorithm for automated driving based on Frenet coordinate system

控制与决策. 2021, 36(4): 815–824 <https://doi.org/10.13195/j.kzyjc.2019.0748>

基于强化学习的多目标车辆跟随决策算法

Multi-objective vehicle following decision algorithm based on reinforcement learning

控制与决策. 2021, 36(10): 2497–2503 <https://doi.org/10.13195/j.kzyjc.2020.0426>

基于概率推理学习优化的无人自行车质量偏心校正方法

黄用华, 梁子彦, 庄 未[†], 任仰华

(桂林电子科技大学 机电工程学院, 广西 桂林 541004)

摘 要: 车体质量偏心是无人自行车一个重要的性能参数, 为了降低车体质量偏心对无人自行车航向轨迹的影响, 提出一种基于有模型强化学习原理的概率推理学习优化 (PILO) 偏心校正方法. 该方法以车体侧向倾角 (倾角速度)、车把转角 (转角速度) 以及车把控制力矩为输入, 以车体侧向倾角速度 (倾角加速度) 以及车把转角速度 (车把转角加速度) 为输出, 利用高斯过程回归 (GPR) 构建系统的概率动态模型 (PDM) 表征系统的不确定性动态, 并将其用于后续的状态序列预测; 将质量偏心作为车把 PD 控制器的一个参数, 考虑车体航向与车把转角间的运动约束, 通过车体航向角速度构造目标函数, 优化并校正系统的质量偏心参数. 设定 8 种不同的负载偏心开展无人自行车仿真以及物理样机实验, 结果表明: PILO 系统校正的绝对误差不超过 0.005 rad, 相对误差低于 10%, 且展现了一定的抗干扰能力; 与无模型的认知学习偏心优化 (RLO) 校正系统相比, PILO 系统在参数整定难度、智能化以及容错能力等方面具有一定优势.

关键词: 无人自行车; 车体航向; 质量偏心校正; 概率推理学习优化; 概率动态模型; 高斯过程回归

中图分类号: TP242

文献标志码: A

DOI: 10.13195/j.kzyjc.2024.0775

引用格式: 黄用华, 梁子彦, 庄未, 等. 基于概率推理学习优化的无人自行车质量偏心校正方法 [J]. 控制与决策, 2025, 40(5): 1723-1732.

A mass eccentricity correction method for unmanned bicycles based on probability inference learning optimization

HUANG Yong-hua, LIANG Zi-yan, ZHUANG Wei[†], REN Yang-hua

(College of Mechanical and Electrical Engineering, Guilin University of Electronic Technology, Guilin 541004, China)

Abstract: Bicycle mass eccentricity is an important performance parameter of unmanned bicycles. In order to reduce the influence of body mass eccentricity on the course trajectory of unmanned bicycles, a probabilistic inference learning optimization (PILO) eccentricity correction method based on the principle of model reinforcement learning is proposed. The method takes the lateral inclination angle (angle velocity), handlebar angle (angle velocity) and handlebar control moment as inputs, and takes the lateral inclination velocity (angle acceleration) and handlebar angle velocity (angle acceleration) of the vehicle body at the corresponding time as outputs. The probabilistic dynamic model (PDM) of the system is constructed by Gaussian process regression (GPR), which is used for the subsequent state sequence prediction. Taking the mass eccentricity as a parameter of the handlebar PD controller, considering the kinematic constraint between the body heading and the handlebar angle, the objective function is constructed by the body heading angular velocity, and the mass eccentricity parameter is optimized and corrected. The results show that the absolute error of the PILO system is less than 0.005 rad, the relative error is less than 10%, and the anti-jamming ability is demonstrated. Compared with the model free cognitive learning eccentric optimization (RLO) calibration system, the PILO system has some advantages in parameter setting difficulty, intelligence and fault tolerance.

Keywords: unmanned bicycle; bicycle heading; mass eccentricity correction; probability inference learning optimization; probability dynamic model; Gaussian process regression

0 引 言

无人自行车是自行车机构与智能控制技术相结

合的产物, 因其轻便、灵活的特点, 在控制理论验证、物流配送和交通出行等领域具有广泛的应用前景.

收稿日期: 2024-07-01; 录用日期: 2024-12-13.

基金项目: 国家自然科学基金项目 (52165001, 51865005, 51765011); 广西自然科学基金项目 (2018GXNSFAA281297, 2018GXNSFAA281301).

责任编辑: 李忠奎.

[†]通信作者. E-mail: zhuangweibupt@sohu.com.

载同样设置为 0, 则行驶时负载会对车轮接地点连线 AB 产生一个顺时针方向的附加力矩使得车体向左倾倒; 2) 此时为了平衡, 车把须左转合成向右的离心力以产生逆时针力矩来抵抗附加力矩避免倾倒; 3) 显然, 上述操作的附带后果是车体航向发生左偏移, 从而导致直线平衡运动受到破坏。

针对上述航向偏移问题, 若将空载时的倾角误差 $e_\varphi = (\varphi - 0)$ 调整为 $e_\varphi = (\varphi - \varphi_b^*)$, 使得综合质心的位置回到车轮接地点正上方, 则理论上可使得无人自行车恢复至空载时的受力状态, 有效减弱质量偏心对于航向跟踪性能的影响, 实现直线行走。可见, 无人自行车的偏心校正问题本质上是寻找合适的偏心期望值 φ_b^* , 使其能够较好地匹配负载偏心带来的控制器倾角误差变化。

2 无人自行车偏心校正方法

2.1 PILO 偏心校正方法

图 3 为构建的无人自行车 PILO 偏心校正系统。图 3 中: 系统底层获取无人自行车的状态数据, 并利用 GPR 学习 PDM 模型; 中层基于该模型和偏心参数执行控制策略进行长期预测, 并以航向角速度构建代价函数来评估策略代价; 顶层通过共轭梯度搜索对控制器偏心参数进行迭代学习, 以实现智能优化。

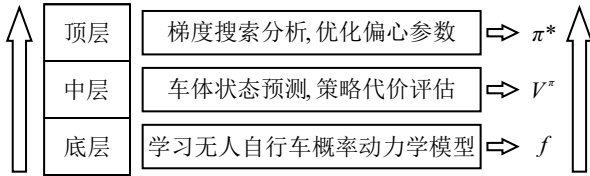


图3 无人自行车 PILO 偏心校正系统结构框架

2.1.1 底层: 学习无人自行车 PDM 模型

PDM 以系统的 5 个状态变量: 横滚角 φ 、横滚角速度 $\dot{\varphi}$ 、车把转角 δ 、车把转角速度 $\dot{\delta}$ 以及车把力矩 τ_δ 为输入向量, 以 4 个状态变量: $\dot{\varphi}$ 、 $\ddot{\varphi}$ 、 $\dot{\delta}$ 以及 $\ddot{\delta}$ 为输出向量, 生成训练样本集。

根据文献 [19], 无人自行车 PDM 模型预测的 4 个变量均值 f_* 以及方差 σ_*^2 计算方法如下所示:

$$\begin{cases} f_* = K_*(K + \sigma_n^2 I)^{-1} y, \\ \sigma_*^2 = K_{**} - K_*(K + \sigma_n^2 I)^{-1} K_*^T. \end{cases} \quad (1)$$

其中: σ_n^2 为噪声方差, n 为训练样本数据长度; $y = [\dot{\varphi}, \ddot{\varphi}, \dot{\delta}, \ddot{\delta}]^T$ 为训练样本输出矢量; K 、 K_* 和 K_{**} 为相应维度的平方指数协方差矩阵; I 为单位矩阵。

采用最大边缘似然的策略来寻求其最优值, 建立的 PDM 模型负对数边际似然函数为

$$-\log p(y|X, \theta) = \frac{1}{2} y^T G^{-1} y + \frac{1}{2} \log |G| + \frac{n}{2} \log 2\pi. \quad (2)$$

其中: $X = [\varphi, \dot{\varphi}, \delta, \dot{\delta}, \tau_\delta]^T$ 为无人自行车模型训练样本输入矢量; $G = K + \sigma_n^2 I$; 超参数 $\theta = [l, \alpha_f, \sigma_n]^T$, 这里 α_f 为信号方差, l 为长度尺度; π 为圆周率。

通过共轭梯度法求解超参数的优化问题 [19], 即可获得无人自行车 PDM 模型。

2.1.2 中层: 车体状态预测以及策略代价评估

利用学习得到的无人自行车 PDM 模型和偏心参数生成倾角误差, 构建控制器并实施控制策略, 对无人自行车进行长期状态推理预测。具体方法如下。

1) 根据当前车体状态 $x_t = [\varphi_t, \dot{\varphi}_t, \delta_t, \dot{\delta}_t]^T$, 计算所对应的控制策略 $\pi(x_t)$ 。

为了避免控制器强鲁棒性对偏心校正效果的干扰, 采用了结构简单的 PD 算法作为控制策略, 即

$$\tau_\delta = k_{P\varphi} e_\varphi + k_{D\varphi} \dot{e}_\varphi + k_{P\delta} e_\delta + k_{D\delta} \dot{e}_\delta. \quad (3)$$

其中: $e_\delta = (\delta - \delta^*)$, δ^* 为 δ 期望值; $k_{P\varphi}$ 、 $k_{D\varphi}$ 、 $k_{P\delta}$ 以及 $k_{D\delta}$ 分别为对应的平衡控制参数。

利用直线轨迹实现车体的质量偏心校正, 因此, 将期望参数 $\dot{\varphi}^*$ 、 δ^* 以及 $\dot{\delta}^*$ 设为 0。令 $\pi(x_t) = \tau_{\delta t}$, 记控制参数矢量 $\Psi = [-k_{P\varphi}, -k_{D\varphi}, -k_{P\delta}, -k_{D\delta}]$ 。为了突出偏心校正问题, 该矢量可事先通过单一的平衡控制 (不考虑航向轨迹) 进行人工整定, 故控制器 (3) 需要优化的参数仅为 $\nu = k_{P\varphi} \varphi^*$, 整理, 有

$$\pi(x_t) = \Psi x_t + \nu. \quad (4)$$

如此, 车体状态与控制策略间的联合分布为

$$p(x_t, \pi(x_t)) = \mathcal{N} \left(\begin{bmatrix} \mu_t \\ \Psi \mu_t + \nu \end{bmatrix}, \begin{bmatrix} \Sigma_t & \Sigma_t \Psi^T \\ \Psi \Sigma_t & \Psi \Sigma_t \Psi^T \end{bmatrix} \right). \quad (5)$$

2) 借助 PDM 模型, 以联合分布 $p(x_t, \pi(x_t))$ 作为模型输入, 推理出车体状态增量 $\dot{x}_t$ 的预测分布为

$$p(\dot{x}_t) \sim \mathcal{N}(\mathbb{E}_{x_t, f}[\dot{x}_t], \text{cov}_{x_t, f}[\dot{x}_t]). \quad (6)$$

对于车体状态预测均值 $\mathbb{E}_{x_t, f}[\dot{x}_t]$ 以及预测协方差矩阵 $\text{cov}_{x_t, f}[\dot{x}_t]$, 对应的数学表达形式如下所示:

$$\mathbb{E}_{x_t, f}[\dot{x}_t] = [\beta_1^T q_1, \beta_2^T q_2, \beta_3^T q_3, \beta_4^T q_4]^T, \quad (7)$$

$$\text{cov}_{x_t, f}[\dot{x}_t] = \begin{bmatrix} \text{var}_{x_t, f}[\dot{\varphi}_t] & \dots & \text{cov}_{x_t, f}[\dot{\varphi}_t, \ddot{\delta}_t] \\ \vdots & \ddots & \vdots \\ \text{cov}_{x_t, f}[\ddot{\delta}_t, \dot{\varphi}_t] & \dots & \text{var}_{x_t, f}[\ddot{\delta}_t] \end{bmatrix}. \quad (8)$$

其中: $\dot{x}_t = [\dot{\varphi}_t, \ddot{\varphi}_t, \dot{\delta}_t, \ddot{\delta}_t]^T$ 为无人自行车的状态预测矢量; $q_a = [q_1, q_2, \dots, q_n]^T$, $a \in \{1, 2, 3, 4\}$ 。

参考文献 [17], 结合已知的模型超参数 θ 、训练样本集输入 X 以及输出 y 、式 (5) 中各元素即可计算

得到车体状态预测均值 $\dot{\mu}_t$ 以及预测协方差矩阵 $\dot{\Sigma}_t$.

3) 根据上述计算得到的车体状态增量预测分布 $p(\dot{x}_t)$, 求得车体的后续状态分布 $p(x_{t+1})$ 为

$$\mu_{t+1} = \mu_t + \mathbb{E}_{x_t, f}[\Delta t \cdot \dot{x}_t], \quad (9)$$

$$\begin{aligned} \Sigma_{t+1} = & \Sigma_t + \text{cov}_{x_t, f}[\dot{x}_t, (\Delta t \cdot \dot{x}_t)] + \\ & \text{cov}_{x_t, f}[(\Delta t \cdot \dot{x}_t), x_t] + \text{cov}_{x_t, f}[\Delta t \cdot \dot{x}_t]. \end{aligned} \quad (10)$$

其中: Δt 为无人自行车系统的伺服周期, 协方差 $\text{cov}_{x_t, f}[\dot{x}_t, (\Delta t \cdot \dot{x}_t)]$ 的计算可参见文献 [17]. 至此, 上述已实现了 $p(x_t)$ 到 $p(x_{t+1})$ 的推理预测, 以此类推, 即可得到 T 步的车体状态分布为

$$p(x_0), p(x_1), p(x_2), \dots, p(x_{T-1}), p(x_T). \quad (11)$$

4) 为了评估质量偏心参数 φ^* 的优化效果, 以航向角速度 $\dot{\psi}_t$ 为评价指标, 构造如下饱和代价函数:

$$c(\dot{\psi}_t) = 1 - \exp \left[-\frac{1}{2}(\dot{\psi}_t - \dot{\psi}^*)^2 \right], \quad (12)$$

其中 $\dot{\psi}^*$ 为车体航向角速度的期望参数, 对于直线轨迹, 此参数设置为 0.

为了将无人自行车的后续状态 x_{t+1} 融入上述代价函数 $c(\dot{\psi}_t)$, 借引文献 [7] 中车体航向角速度与车把转角 (速度) 间的线性映射关系, 即

$$\dot{\psi}_t = \frac{v}{\omega} \sin(\alpha) \delta_t + \frac{c}{\omega} \sin(\alpha) \dot{\delta}_t. \quad (13)$$

其中: v 为车体前进速度, ω 为车体前后轮间的轴距, α 为车叉前倾角, c 为前轮接地点的后偏距. 最后, 整合每一时刻的代价 $c(\dot{\psi}_t)$, 可得到总代价

$$V^{\pi_{\varphi^*}}(\dot{\psi}_0) = \sum_{t=0}^T \mathbb{E}_{\dot{\psi}_t} [c(\dot{\psi}_t)] = \sum_{t=0}^T \int c(\dot{\psi}_t) p(\dot{\psi}_t) d\dot{\psi}_t. \quad (14)$$

2.1.3 顶层: 梯度搜索分析以及优化偏心参数

构造以下优化模型:

$$\varphi^* \in \arg \min_{\varphi^* \in \Pi} V^{\pi_{\varphi^*}}(\dot{\psi}_0), \quad (15)$$

其中 Π 为偏心参数 φ^* 的备选策略空间. 采用梯度搜索方法对式 (15) 描述的优化问题进行求解. 策略总代价函数对偏心参数的梯度如下所示:

$$\frac{dV^{\pi_{\varphi^*}}(\dot{\psi}_0)}{d\varphi^*} = \sum_{t=0}^T \frac{d}{d\varphi^*} \mathbb{E}_{\dot{\psi}_t} [c(\dot{\psi}_t)]. \quad (16)$$

由于 $\dot{\psi}_t$ 服从高斯分布, 式 (16) 右项变为

$$\begin{aligned} \frac{d}{d\varphi^*} \mathbb{E}_{\dot{\psi}_t} [c(\dot{\psi}_t)] = & \frac{\partial \mathbb{E}_{\dot{\psi}_t} [c(\dot{\psi}_t)]}{\partial \mathbb{E}_{\dot{\psi}_t, \dot{\delta}_t} [\dot{\psi}_t]} \cdot \frac{d \mathbb{E}_{\dot{\psi}_t, \dot{\delta}_t} [\dot{\psi}_t]}{d\varphi^*} + \\ & \frac{\partial \mathbb{E}_{\dot{\psi}_t} [c(\dot{\psi}_t)]}{\partial \text{cov}_{\dot{\psi}_t, \dot{\delta}_t} [\dot{\psi}_t]} \cdot \frac{d \text{cov}_{\dot{\psi}_t, \dot{\delta}_t} [\dot{\psi}_t]}{d\varphi^*}. \end{aligned} \quad (17)$$

理论上, 经过迭代可得到偏心参数阶段性的最

优值. 循环重复第 2.1.1 节 ~ 第 2.1.3 节的学习过程, 当车体实现直线轨迹跟踪 (车体航向角速度稳定在 0 值附近) 时, 即可认为已实现车体质量偏心的校正.

2.2 RLO 偏心校正方法

为了对标 PILO 偏心校正系统的性能, 设计无人自行车认知学习优化 (RLO) 偏心校正系统^[20], 如图 4 所示.

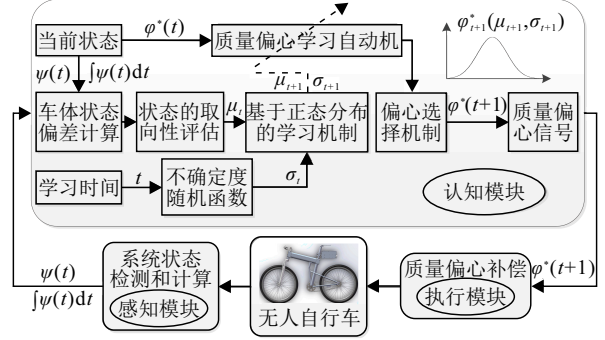


图4 无人自行车 RLO 偏心校正系统

RLO 偏心校正系统由感知、认知和执行 3 个模块组成. 首先, 感知模块实时获取无人自行车的航向角及其积分值, 将数据传递给认知模块; 然后, 认知模块对航向信息进行推理和评估, 更新自动机组, 并将通过偏心选择机制输出优化后的质量偏心参数传递至执行模块; 最后, 执行模块补偿该参数到平衡控制器来生成力矩信号驱动车把电机, 以此调整车体的横滚姿态. 通过这一过程, 无人自行车与外界持续进行“交互-试错”, 不断校正车体的质量偏心.

与 PILO 偏心校正系统不同, RLO 偏心校正系统在学习过程中不依赖具体的系统模型, 其核心在于比对当前时刻与上一时刻的状态评价值, 计算系统状态取向函数, 以此明确偏心参数的学习方向.

3 数值仿真

3.1 PILO 偏心校正

以 Matlab/Simulink 为仿真平台, 搭建无人自行车 PILO 偏心校正数值仿真系统, 如图 5 所示.

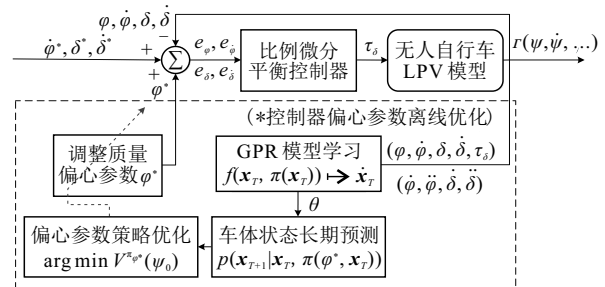


图5 PILO 偏心校正仿真系统

以文献 [21] 中给出的经典无人自行车线性变参数 (LPV) 力学模型为仿真对象, 其中构建 LPV 模型

所需的物理参数名义值可通过虚拟样机和实体物理样机测得, 具体如表 1 和表 2 所示。

表1 无人自行车整体物理参数

参数	数值
车体的总质量 m_t/kg	14.85
前后轮之间的轴距 ω/m	0.885
车体质心高度 h/m	0.547
车体质心到坐标系原点的水平距离 b/m	0.344
前轮接地点的后偏距 c/m	0.078
车叉前倾角 $\alpha/(\circ)$	75

表2 无人自行车构件的物理参数

物理参数	无人自行车部件			
	后轮(R)	车架(T)	前叉(H)	前轮(F)
质量 m/kg	1.713	8.989	2.435	1.713
$J_{xx}/(\text{kg}\cdot\text{m}^2)$	0.056	0.449	0.091	0.056
$J_{yy}/(\text{kg}\cdot\text{m}^2)$	0.111	0.848	0.072	0.111
$J_{zz}/(\text{kg}\cdot\text{m}^2)$	0.056	0.430	0.029	0.056
$J_{xz}/(\text{kg}\cdot\text{m}^2)$	0.000	-0.073	0.016	0.000
质心 $(x, z)/\text{m}$	(0.000, -0.293)	(0.299, -0.521)	(0.816, -0.579)	(0.885, -0.293)

注: J_{xx} (J_{yy} J_{zz} J_{xz}) 为惯性张量。

设置仿真时长为 60 s, 仿真步长为 5 ms; 假设车体为匀速行驶, 车速为 1.8 m/s; 车体横滚角 (速度)、车把转角 (速度) 初始值均设定为 0 rad (rad/s)。另外, 在仿真过程中注入多种不同类型的干扰信号, 具体如下。

1) 在第 5 s、第 15 s 分别注入幅值为 2 N·m 和 -2 N·m、宽度均为 1 s 的车架侧向脉冲力矩干扰; 在第 10 s、第 20 s 分别注入幅值为 2 N·m 和 -2 N·m、宽度均为 1 s 的车把脉冲力矩干扰。

2) 在仿真全程注入幅值为 2 N·m 的随机车把力矩干扰和随机侧向力矩干扰, 并给车体速度也注入幅值为 0.25 m/s 的随机干扰。

3) 在仿真全程给车体横滚角、横滚角速度、车把转角以及转角速度 4 种状态的测量传感器注入幅值为 0.01 rad (rad/s) 的随机干扰。

结合经验分析车体 PD 控制的灵敏性和收敛范围等特性, 通过人工反复调试, 设定控制器参数为 $k_{P\varphi} = -28.07$; $k_{D\varphi} = -14.96$, $k_{P\delta} = -0.94$; $k_{D\delta} = -0.46$ 。状态对应期望值 δ^* 、 $\dot{\varphi}^*$ 、 $\dot{\delta}^*$ 分别为 0 rad 或 0 rad/s。

另外, 在 PILO 偏心校正模块的长期预测中, 假设车体初始时刻为直立状态, 即设定状态均值 μ_0 为 0 rad (rad/s); 考虑到实际应用中, 车体的真实状态

值受传感器精度的影响, 因此, 设定状态方差 Σ_0 为对角元素均为 0.01 的 4×4 阶对角矩阵; 设定初始偏心参数 $\varphi^* = 0 \text{ rad}$, 预测步数为 67; 设定当车体航向角的绝对值 $|\Psi_t| > 1 \text{ rad}$ 时 (对应为车体倾倒的趋势), 终止当前学习, 并保存上一轮校正后得到的偏心参数 φ^* 作为下一轮优化校正的初始策略参数。

设置如表 3 所示的 4 种负载场景进行仿真。

表3 无人自行车负载质量以及位置参数

仿真场景	负载质量/kg	负载位置	偏心的名义值/rad
场景1	3	(0, -0.25, 0.65)	-0.0745
场景2	6	(0, -0.25, 0.65)	-0.1248
场景3	3	(0, -0.25, 0.45)	-0.0792
场景4	3	(0, 0.45, 0.65)	0.1340

以场景 1 为例, 经 4 轮学习即可完成偏心参数的优化校正, 每轮学习的系统仿真数据如表 4 所示。

表4 场景 1 中各轮学习的系统仿真数据

评估参数	初始	第1轮	第2轮	第3轮	第4轮
学习时长/s	2.5	4.2	6.2	29	60
偏心值/rad	0	0.0351	0.0601	0.0822	0.0739

由表 4 可知, 4 轮优化后的偏心校正结果为 0.073 9 rad, 与其名义值 0.074 5 rad (表 3) 相差无几。

考察 4 轮 PILO 偏心校正后得到的车体航向角状态, 对应的状态曲线如图 6 所示。

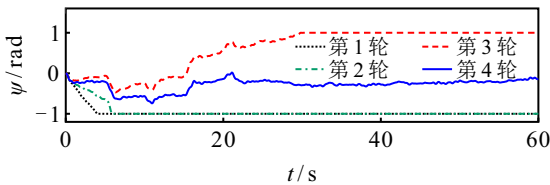


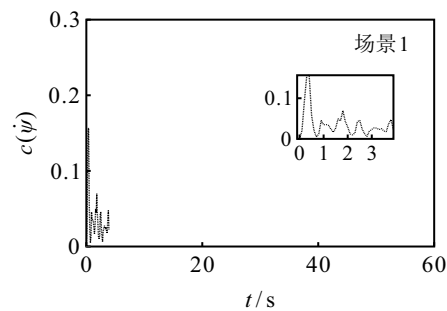
图6 车体航向角状态数据曲线 (场景 1)

由图 6 可见, 车体在第 4 轮学习后稳定收敛于 $[-0.265\ 5, -0.175\ 6](\text{rad})$ 范围内, 表明车体基本实现直线轨迹运动, 偏心校正有效。

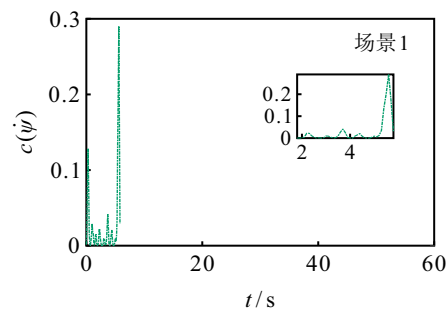
进一步地, 取每轮学习的有效状态数据代入式 (12), 绘制代价函数曲线, 如图 7 所示。由图 7 可见, 随着优化深入, 航向代价值逐渐收敛至 0 值附近 (如图 7(d) 所示), 表明 PILO 偏心校正有效。

图 8 为其余场景最终的航向角状态曲线。由图 8 可见, 3 种负载场景下无人自行车系统的车体航向角在 26 s 后基本实现直线轨迹行驶。

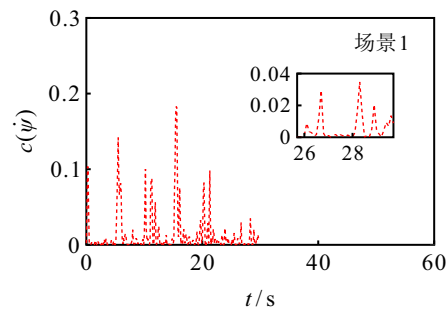
整合上述 4 种场景的仿真数据, 如表 5 所示。由表 5 的仿真数据可见: 1) 随着该系统的迭代学习优化, 4 种不同场景的偏心参数 φ^* 逐渐向对应的车体



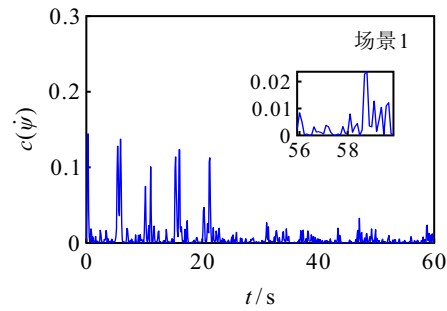
(a) 第1轮优化结果



(b) 第2轮优化结果



(c) 第3轮优化结果



(d) 第4轮优化结果

图7 4轮优化学习的系统仿真代价曲线(场景1)

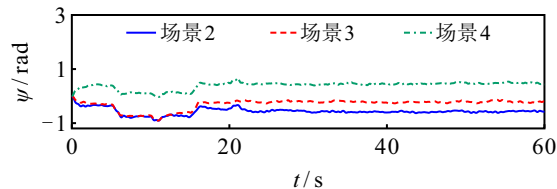


图8 其余3种不同场景下航向角状态数值仿真曲线

质量偏心优化目标,即质量偏心名义值的相反数 $-\varphi_b^*$ 收敛,而上述情形与第2.2节的机理分析相符;2)在4种不同场景下,偏心校正的绝对误差均维持在0.002rad,相对误差不超过2%,表明该PILO偏心校正精度满足控制需求;3)即便对系统施加多源干

表5 4种不同场景的PILO仿真数据

仿真场景	优化系统评估参数			
	优化结果/rad	优化目标/rad	绝对误差/rad	相对误差/%
场景1	0.0739	0.0745	0.0006	0.81
场景2	0.1235	0.1248	0.0013	1.04
场景3	0.0798	0.0792	0.0006	0.76
场景4	-0.1332	-0.1340	0.0008	0.61

扰,其对PILO偏心校正的结果影响有限,表明本节所设计的优化系统具备一定的抗干扰能力。

3.2 RLO 偏心校正

类似地,以Matlab/Simulink为仿真平台,搭建RLO偏心校正仿真系统.配置与PILO相同的仿真环境,仿真结果如图9所示.由图9可见,RLO同样能够使得系统收敛.综合4种场景分析RLO的校正精度,具体数据如表6所示。

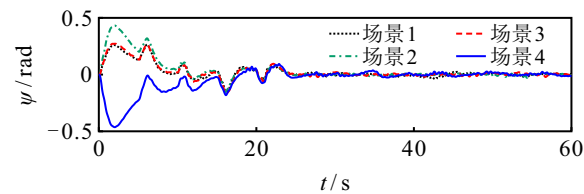


图9 RLO系统航向角状态数据曲线(场景1)

表6 4种不同场景的RLO仿真数据

仿真场景	优化系统评估参数			
	优化结果/rad	优化目标/rad	绝对误差/rad	相对误差/%
场景1	0.0742	0.0745	0.0003	0.40
场景2	0.1242	0.1248	0.0006	0.48
场景3	0.0785	0.0792	0.0007	0.88
场景4	-0.1345	-0.1340	0.0005	0.37

由图9和表6可见,RLO偏心校正后的车体航向角稳定在一定范围内,表明了该系统的有效性。

3.3 两种偏心校正结果对比分析

将表5与表6的数据进行对比,结果如下:1)两种校正系统的精度相似,偏心参数的绝对误差均在0.001rad附近,相对误差约为1%。2)PILO系统的平均绝对误差为0.00083rad,平均相对误差为0.805%;而RLO系统的对应值为0.00053rad和0.533%。这表明RLO系统的校正精度略优,其原因主要在于其融入了先验知识.RLO系统的偏心选择机制具有随机性,导致优化过程不确定,因此,在仿真中加入了小权重的先验指导以提高收敛性能,这也表明PILO系统在智能化程度上更高.另外,仿真中发现,PILO的学习参数较少,调试时对优化效率和精度不敏感;而RLO包含更多学习参数,大多数对学习效率和优化精度敏感,调试过程较为繁琐和耗时.究其动因是经验空间不同,PILO方法因拥有更多经验指导,在

容错率上更具优势.

综上所述, 尽管两种系统的校正精度相似, 但是 PILO 系统在智能化程度和容错率上优于 RLO. 此外, PILO 偏心校正系统在调试阶段无需频繁调整学习参数, 展现出更好的普适性.

4 物理样机实验

实验用到的物理样机如图 10 所示.

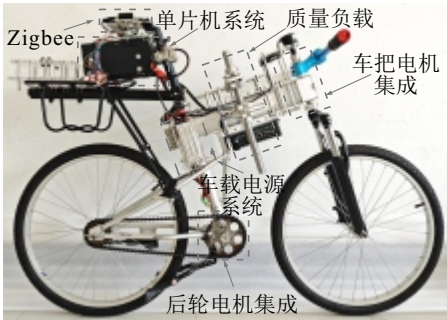


图10 无人自行车物理样机

物理样机上装有单片机系统、ZigBee 模块、车载电源系统等, 两个电机集成模块分别安装在车把和脚蹬的关节内部来驱动车轮和车把, 其中无人自行车样机测控系统的具体架构如图 11 所示.

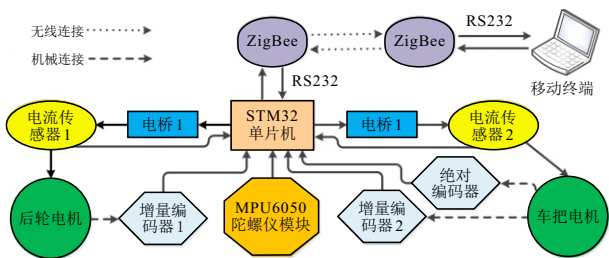


图11 无人自行车物理样机测控系统框架

图 11 中: 增量编码器和绝对编码器用于测量车体后轮转速和车把转角; MPU6050 陀螺仪测量车体姿态, 并产生 5 ms 的时基作为系统伺服的节拍; 电流传感器 1 和电流传感器 2 分别采集后轮电机和车把电机的驱动电流; ZigBee 负责将车体状态数据上传至移动终端.

与数值仿真的场景设计类似, 样机实验共设置 8 种不同的负载场景展开实施, 如表 7 所示. 表 7 中的高度(Z)和偏心距(Y)定义见图 2.

4.1 PILO 偏心校正系统样机实验

物理样机与 LPV 力学模型存在结构偏差, 且真实工况 (如地形、路况和驱动信号, 特别是样机的电机驱动信号为 PWM 调制, 而仿真为力矩) 与仿真工况差异较大, 因此, 需要重新选择控制参数, 具体如下.

结合经验分析无人自行车 PD 控制的灵敏性和收敛范围, 通过人工反复调试, 设定 PD 平衡控制器

表7 物理样机实验方案参数

实验方案	加载位置和质量		
	高度(Z)/m	偏心距(Y)/m	质量(Q)/kg
方案1	-0.65	-0.15	3.10
方案2	-0.65	0.15	3.10
方案3	-0.65	-0.15	4.36
方案4	-0.65	0.15	4.36
方案5	-0.55	-0.15	3.10
方案6	-0.55	0.15	3.10
方案7	-0.65	-0.30	3.10
方案8	-0.65	0.30	3.10

参数为 $k_{P\varphi} = -27\,680.84$ 、 $k_{D\varphi} = -75\,073.2$ 、 $k_{P\delta} = -2.15$ 、 $k_{D\delta} = -0.46$; 车体速度设定为 2 m/s; 考虑外界扰动对传感器精度的影响, 调整初始方差 Σ_0 的对角元素为 0.02; 设定当车体航向角绝对值 $|\Psi_t| > 0.3\text{ rad}$ 时 (对应为车体倾倒趋势) 终止学习; 其余参数与仿真一致.

图 12 为方案 1 和方案 8 的加载细节. 实验前先通过较为繁杂的人为勘探, 获取车体偏心估计值: $\phi_{\text{方案1}}^* = -0.039\text{ rad}$, $\phi_{\text{方案8}}^* = 0.068\text{ rad}$.

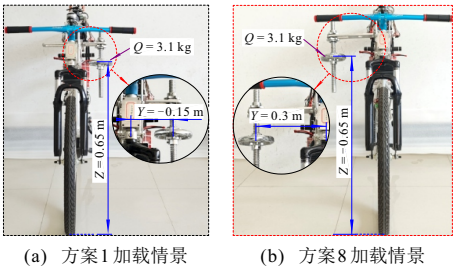


图12 无人自行车物理样机实验加载情景

样机实验在户外水平路面展开. 图 13 为方案 1 和方案 8 偏心校正后自行车运动实验视频连拍图. 由图 13 可见, 自行车在方案 1 和方案 8 中均能够依次通过标有数字的路障 (0, 1, ..., 路障间距为 2.1 m), 且两车轮可在相距 0.8 m 的黑色胶带之间行驶.



(a) 方案1 样机实验视频连拍图



(b) 方案8 样机实验视频连拍图

图13 PILO 偏心校正后的无人自行车平衡运动实验视频

进一步给出方案 1 中各轮次优化后的车体航向角数据, 如图 14 所示. 由图 14 方案 1 中的数据曲线可见, 偏心参数尚未优化时 ($\varphi^* = 0\text{ rad}$), 无人自行

车受质量偏心的影响,车体航向轨迹向一侧出现大幅度的偏移,最终在 2.7 s 附近车体航向角到达系统所设定的临界值($\Psi = -0.3 \text{ rad}$),停止本轮学习。

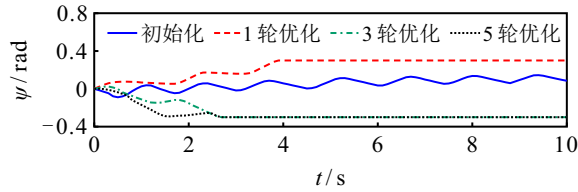


图14 PILO 偏心校正进程实验数据曲线 (方案1)

在经过5轮偏心参数优化后,车体航向角 Ψ 稳定在 $[-0.0729, 0.1447]$ (rad)的区间范围内,无人自行车基本实现直线轨迹行驶。以最后一轮优化的质量偏心参数 $\varphi^* = 0.042 \text{ rad}$ 作为方案1最终的校正值,其与偏心估计值的相反数 $-\phi_{\text{方案1}}^* = 0.039 \text{ rad}$ 的绝对误差为0.003 rad,相对误差为7.69%。

图15为方案8各轮次的车体航向角数据。与方案1不同,方案8中车体负载的悬挂位置相反,且偏心程度更大。因此,由图15的航向角数据曲线可见,在偏心参数未优化时($\varphi^* = 0 \text{ rad}$),无人自行车受到质量负载影响,航向轨迹向相反一侧偏移更大,直至0.8 s附近才达到系统设定的临界值($\Psi = 0.3 \text{ rad}$),停止本轮学习。即便如此,经5轮迭代,车体航向角仍然能够稳定在 $[-0.1226, 0.0569]$ (rad)范围内,基本实现直线行驶,此时偏心参数 $\varphi^* = -0.070 \text{ rad}$,与 $-\phi_{\text{方案8}}^* = -0.068 \text{ rad}$ 的绝对误差为0.002 rad,相对误差为2.94%。

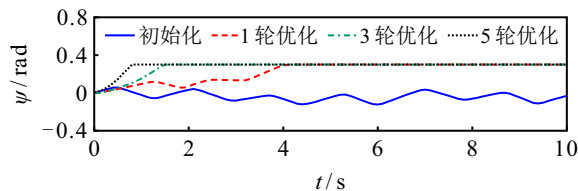


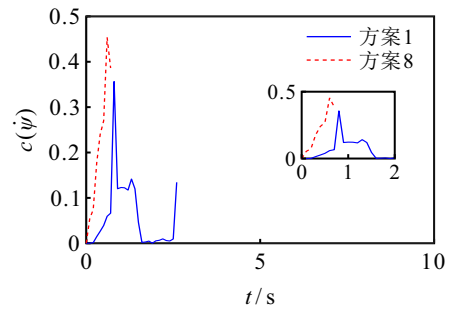
图15 PILO 偏心校正进程实验数据曲线 (方案8)

为了更直观地分析PILO偏心校正系统的优化性能,对上述方案1和方案8中系统各轮学习优化的样机实验数据进行整合梳理,具体如表8所示。表8表明,两种方案的偏心参数 φ^* 逐轮向各自对应的 $-\phi_{\text{方案1}}^*$ (0.039 rad)和 $-\phi_{\text{方案8}}^*$ (-0.068 rad)逼近。

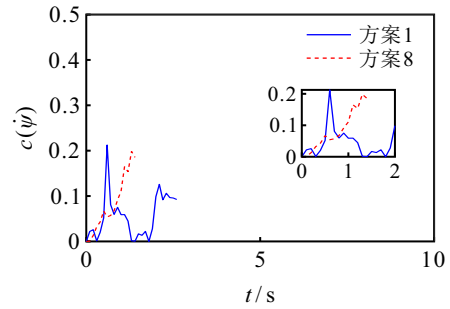
表8 各轮次优化的偏心参数实验数据 (方案1和方案8)

实验方案	控制器偏心参数 φ^*/rad				
	1轮优化	2轮优化	3轮优化	4轮优化	5轮优化
方案1	0.014	0.030	0.058	0.046	0.042
方案8	-0.030	-0.038	-0.046	-0.071	-0.070

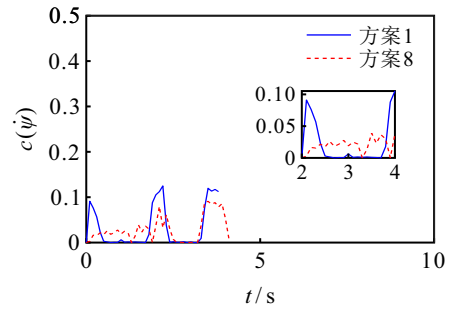
进一步地,考察这两种方案优化前、1轮、3轮以及5轮优化后系统代价的演化,如图16所示。由图16可见,经过PILO的5轮优化,自行车系统的代价



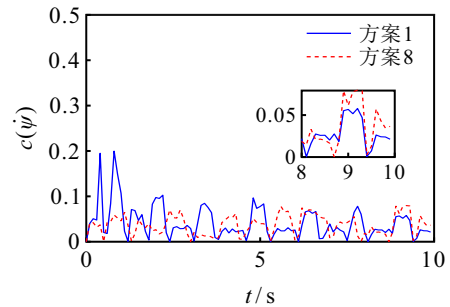
(a) 未开始偏心参数优化



(b) 第1轮偏心参数优化



(c) 第3轮偏心参数优化



(d) 第5轮偏心参数优化

图16 PILO 偏心校正实验代价曲线 (方案1和方案8)

曲线整体收敛在0值附近。

综上初步有结论:PILO系统在工程中可有效实现偏心参数校正,使得车体基本实现直线轨迹行驶。

同样地,将上述8种方案的数据整合为表9。由表9可知,经4轮~6轮迭代,所有场景偏心参数的绝对误差均在0.005 rad以内,相对误差低于10%。另外,偏心校正后的车体航向角稳定在小范围内,校正精度满足控制需求。这与数值仿真结果一致。

综上分析可以表明:对于不同的负载条件,PILO偏心校正系统均可实现车体质量偏心的优化校正,具有一定的可行性、可靠性。

表9 8种方案下 PILO 偏心校正系统实验数据

实验方案	优化次数	系统优化值/rad	测量相反值/rad	绝对误差/rad	相对误差/%
方案1	5	0.042	0.039	0.003	7.69
方案2	4	-0.039	-0.037	0.002	5.41
方案3	5	0.043	0.042	0.001	2.38
方案4	6	-0.045	-0.044	0.001	2.27
方案5	4	0.058	0.060	0.002	3.33
方案6	4	-0.061	-0.059	0.002	3.39
方案7	6	0.070	0.071	0.001	1.41
方案8	5	-0.070	-0.068	0.002	2.94

表10 8种方案下 RLO 偏心校正系统实验数据

实验方案	优化次数	系统优化值/rad	测量相反值/rad	绝对误差/rad	相对误差/%
方案1	2	0.037	0.039	0.002	5.13
方案2	2	-0.034	-0.037	0.003	8.11
方案3	3	0.040	0.042	0.002	4.76
方案4	3	-0.046	-0.044	0.002	4.55
方案5	2	0.058	0.060	0.002	3.33
方案6	2	-0.058	-0.059	0.001	1.69
方案7	3	0.073	0.071	0.002	2.82
方案8	3	-0.063	-0.068	0.005	7.35

4.2 RLO 偏心校正系统样机实验

RLO 偏心校正实验与 PILO 一样在 8 种不同负载场景下进行, 同样以方案 1 和方案 8 为典型案例, 实验视频如图 17 所示。



(a) 方案1 样机实验视频连拍图



(b) 方案8 样机实验视频连拍图

图17 RLO 偏心校正后车体平衡运动实验视频

由图 17 可见, 无人自行车在方案 1 和方案 8 两种不同质量负载下的行驶状态与第 4.1 节的效果相似。因此, RLO 系统同样能够有效校正这两种负载下的质量偏心。

进一步给出 RLO 偏心校正系统作用下, 方案 1 和方案 8 最后一轮优化的车体航向曲线如图 18 所示。由图 18 可知, 对于方案 1 和方案 8, 车体航向角 ψ 最终稳定在 $[-0.039, 0.105](\text{rad})$ 和 $[-0.148, 0.099](\text{rad})$ 区间范围内, 无人自行车均基本实现直线行驶。

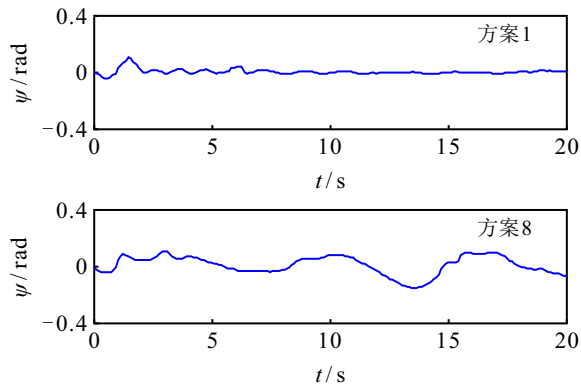


图18 偏心校正后车体平衡运动航向曲线
(方案 1 和方案 8)

表 10 为综合方案 1 ~ 方案 8 的实验数据结果。

分析表 10 的实验数据可知, 在 8 种不同负载方案下, 经过 2 轮、3 轮迭代优化, RLO 偏心校正系统成功校正了无人自行车的质量偏心, 且校正精度同样满足控制需求, 该结论与数值仿真结果一致。

综上可得出结论, RLO 偏心校正系统在工程应用上同样具有一定的可行性和可靠性。

4.3 两种偏心校正系统比对分析

对比表 9 与表 10 两种校正系统的数据可知: 1) 两种系统的校正绝对误差均维持在 0.005 rad 范围内, 相对误差在 10 % 以内, 表明两种偏心校正系统的校正精度相似。2) PILO 系统的平均绝对误差为 0.001 75 rad, 平均相对误差为 3.602 %; 而 RLO 系统分别为 0.002 38 rad 和 4.718 %。与仿真结果不同, PILO 系统在质量偏心校正精度上优于 RLO 系统, 原因在于 RLO 系统在干扰下偏心参数的收敛受到影响, 降低了识别精度。现实中, 侧向风扰和路面不平不确定性干扰会进一步削弱 RLO 的校正性能, 而 PILO 系统通过持续学习和经验积累, 更能适应外部扰动, 使得结果更准确、稳定。此外, 调试过程中发现, RLO 系统的学习参数对于效率和精度更敏感, 需要多次繁琐地调试, 且收敛速度与初始值密切相关, 降低了容错率。

综上所述, PILO 偏心校正系统在实际应用中智能化程度更高, 具有更好的鲁棒性和适应能力。

5 结 论

本文以无机械平衡调节器的无人自行车为对象, 基于有模型强化学习的 PILO 方法构建了一种质量偏心校正系统, 探讨了其在不同负载状态下的校正问题。

1) 通过分析质量偏心对车体航向轨迹的影响, 结合传统 PD 控制器, 将质量偏心作为车把控制输入的重要性能参数, 可快速有效地建立无人自行车的质量偏心校正系统。

2) 基于概率推理学习优化方法设计的 PILO 偏心校正系统, 能够在不同负载下准确校正车体质量

偏心, 实现直线轨迹, 并具备可靠性和抗干扰能力; 相比于 RLO 系统, PILO 在智能化程度和适用性上具有优势。

3) 无人自行车在不同负载下的质量偏心校正, 可减小其对于航向跟踪性能的影响, 能够为后续的航向轨迹跟踪研究提供理论支持和技术参考。

参考文献 (References)

- [1] Kooijman J D G, Meijaard J P, Papadopoulos J M, et al. A bicycle can be self-stable without gyroscopic or caster effects[J]. *Science*, 2011, 332(6027): 339-342.
- [2] Alizadeh M, Ramezani A, Saadatinezhad H. Fault tolerant control in an unmanned bicycle robot via sliding mode theory[J]. *IET Cyber-Systems and Robotics*, 2022, 4(2): 139-152.
- [3] Zhao H, Liu W Y, Chen X L. Adaptive robust constraint-following control for underactuated unmanned bicycle robot with uncertainties[J]. *ISA Transactions*, 2023, 143: 144-155.
- [4] Huang Y H, Liao Q Z, Guo L, et al. Simple realization of balanced motions under different speeds for a mechanical regulator-free bicycle robot[J]. *Robotica*, 2015, 33(9): 1958-1972.
- [5] Huang Y H, Liao Q Z, Guo L, et al. Balanced motions realization for a mechanical regulators free and front-wheel drive bicycle robot under zero forward speed[J]. *International Journal of Advanced Robotic Systems*, 2013, 10(317): 1-9.
- [6] Zhang Y L, Liu Y, Yi G R. Model analysis of unmanned bicycle and variable gain LQR control[C]. *Proceedings of the IEEE 9th Joint International Information Technology and Artificial Intelligence Conference*. Chongqing, 2020: 1196-1200.
- [7] Baquero-Suárez M, Cortés-Romero J, Arcos-Legarda J, et al. A robust two-stage active disturbance rejection control for the stabilization of a riderless bicycle[J]. *Multibody System Dynamics*, 2019, 45(1): 7-35.
- [8] Hashemnia S, Shariat P M, Mahjoob M J. Unmanned bicycle balancing via Lyapunov rule-based fuzzy control[J]. *Multibody System Dynamics*, 2014, 31: 147-168.
- [9] 李艳, 王涵. 基于复合控制的自行车机器人平衡控制[J]. *现代电子技术*, 2019, 42(2): 177-182.
(Li Y, Wang H. Balance control of bicycle robot based on compound control[J]. *Modern Electronics Technique*, 2019, 42(2): 177-182.)
- [10] Fang Z X, Savkin A V. Strategies for optimized UAV surveillance in various tasks and scenarios: A review[J]. *Drones*, 2024, 8(5): 193.
- [11] Jin L S, Zhou H P, Xie X Y, et al. A direct yaw moment control frame through model predictive control considering vehicle trajectory tracking performance and handling stability for autonomous driving[J]. *Control Engineering Practice*, 2024, 148: 105947.
- [12] Xiao H Z, Chen C H, Zhang G D, et al. Reinforcement learning-driven dynamic obstacle avoidance for mobile robot trajectory tracking[J]. *Knowledge-Based Systems*, 2024, 297: 111974.
- [13] Miranda-Colorado R, Cazarez-Castro N R. Observer-based fuzzy trajectory-tracking controller for wheeled mobile robots with kinematic disturbances[J]. *Engineering Applications of Artificial Intelligence*, 2024, 133: 108279.
- [14] Chen C, Zhang X L, Peng X F. Trajectory tracking control of four-rotor UAV based on nonlinear extended state observer and model predictive control in wind disturbance environment[J]. *Journal of Physics: Conference Series*, 2024, 2764(1): 012075.
- [15] 王若琳, 胡翔, 余焯, 等. 基于扭摆的落体质心与光心距离的高精度测量[J]. *仪器仪表学报*, 2022, 43(8): 85-92.
(Wang R L, Hu X, Yu Y, et al. High-precision measurement of the distance between the mass center and the optical center of falling body based on the pendulum system[J]. *Chinese Journal of Scientific Instrument*, 2022, 43(8): 85-92.)
- [16] 郭亚内. 多轴车辆质心测量方法分析[J]. *黑龙江科学*, 2019, 10(14): 70-71.
(Guo Y N. Analysis of centroid measurement method for multi-axle vehicles[J]. *Heilongjiang Science*, 2019, 10(14): 70-71.)
- [17] Deisenroth M P, Rasmussen C E. PILCO: A model-based and data-efficient approach to policy search[C]. *Proceedings of the 28th International Conference on Machine Learning*. Madison, 2011: 465-472.
- [18] Feinberg V, Wan A, Stoica I, et al. Model-based value estimation for efficient model-free reinforcement learning[J/OL]. 2018, arXiv: 1803.00101.
- [19] 苏国韶, 张研, 肖义龙. 高斯过程机器学习及其工程应用[M]. 北京: 科学出版社, 2020.
(Su G S, Zhang Y, Xiao Y L. Gaussian process for machine learning and its applications in engineering[M]. Beijing: Science Press, 2020.)
- [20] 赵传松, 任红格, 史涛, 等. 内在动机轮式倒立摆反应式认知系统[J]. *浙江大学学报: 工学版*, 2018, 52(6): 1073-1080.
(Zhao C S, Ren H G, Shi T, et al. Wheeled inverted pendulum reactive cognitive system with internal motivation[J]. *Journal of Zhejiang University: Engineering Science*, 2018, 52(6): 1073-1080.)
- [21] Kooijman J D G, Schwab A L, Meijaard J P. Experimental validation of a model of an uncontrolled bicycle[J]. *Multibody System Dynamics*, 2008, 19(1): 115-132.

作者简介

黄用华 (1977-), 男, 研究员, 博士, 主要研究方向为机器人技术、机器人动力学与控制、机器视觉与认知计算, E-mail: huangyonghuaxj@sina.com;

梁子彦 (1999-), 男, 硕士生, 主要研究方向为机器人动力学与控制、机器学习与认知计算, E-mail: 827123241@qq.com;

庄未 (1977-), 女, 研究员, 博士, 主要研究方向为欠驱动机器人技术、机器人动力学与控制、嵌入式测控系统, E-mail: zhuangweibupt@sohu.com;

任仰华 (1999-), 女, 硕士生, 主要研究方向为机器人技术、机器人动力学与控制, E-mail: 1318898422@qq.com.