

虚拟无人机实验 1-第 1 组

实验名称: Unity AirSim + APF + DDPG 权重优化仿真训练

实验描述: 使用 Unity AirSim 仿真环境，结合人工势场算法（APF）和深度确定性策略梯度算法（DDPG），通过强化学习训练优化 APF 算法的 5 个权重系数（ $\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5$ ），实现多无人机协同探索任务的智能权重分配。

输入:

- **仿真环境:** Unity AirSim 多无人机 3D 仿真环境
- **无人机配置:** 4 台无人机协同训练（UAV1, UAV2, UAV3, UAV4）
 - 主训练无人机: UAV1（用于 DDPG 学习）
 - 协同无人机: UAV2, UAV3, UAV4（提供环境交互）
- **状态空间:** 18 维连续状态空间
 - 位置信息（3 维）: [x, y, z]
 - 速度信息（3 维）: [vx, vy, vz]
 - 方向信息（3 维）: [dx, dy, dz]
 - 熵值信息（3 维）: [平均熵值, 最大熵值, 已扫描比例]
 - Leader 信息（3 维）: [到 Leader 距离, 相对位置 x, 相对位置 z]
 - 扫描信息（3 维）: [扫描单元数, 扫描增量, 扫描趋势]
- **动作空间:** 5 维连续动作空间（APF 权重系数）
 - α_1 : repulsionCoefficient（排斥力权重）
 - α_2 : entropyCoefficient（信息熵权重）
 - α_3 : distanceCoefficient（就近权重）
 - α_4 : leaderRangeCoefficient（边界力权重）
 - α_5 : directionRetentionCoefficient（保持方向权重）
- **训练参数:**
 - 总训练步数: 100 步
 - 每步飞行时长: 20.0 秒
 - 检查点保存频率: 每 1000 步
 - 学习率: $1e-4$
 - 经验回放缓冲区大小: 5000
 - 批次大小: 64
 - 折扣因子 γ : 0.99
 - 软更新系数 τ : 0.005
 - 动作噪声: 正态分布噪声（均值 0, 标准差 0.15）

输出:

- **训练模型文件:**
 - weight_predictor_airsim.zip - 最终训练完成的 DDPG 模型（用于 APF 权重预测）
 - best_model.zip - 训练过程中最佳奖励性能的模式
 - checkpoint_*.zip - 训练过程中的检查点模型文件（每 1000 步保存一次）

- **模型功能:** 训练好的模型可以根据当前环境状态 (18 维) 预测最优的 APF 权重系数 (5 维), 用于指导多无人机协同探索任务中的路径规划和避障行为

实验流程:

1. 启动 Unity AirSim 仿真环境 (./start.bat)
2. 初始化 AlgorithmServer, 连接 4 台无人机
3. 启动无人机任务和算法线程
4. 创建 SimpleWeightEnv 训练环境
5. 初始化 DDPG 模型 (Actor-Critic 架构)
6. 执行训练循环:
 - 获取当前环境状态 (18 维)
 - DDPG 模型预测权重系数 (5 维)
 - 将权重应用到 APF 算法
 - 执行动作并观察奖励
 - 更新 DDPG 模型参数
7. 保存训练完成的模型
8. 清理资源并结束训练

关键技术:

- 强化学习算法: DDPG (Deep Deterministic Policy Gradient)
- 路径规划算法: APF (Artificial Potential Field)
- 仿真平台: Unity AirSim
- 训练框架: Stable-Baselines3