

## 新闻中心

---

[科研动态 \(../\)](#)

[近日要闻 \(../jryw/\)](#)

[媒体扫描 \(../mtsm/\)](#)

[头条新闻 \(../ttxw/\)](#)

[学术活动 \(../xshd/\)](#)

[产业化动态 \(../cyhdt/\)](#)

[信息公开 \(../xxgk/\)](#)

当前位置: [首页 \(../../\)](#) > [新闻中心 \(../..\)](#) > [科研动态 \(../\)](#)

**科研动态**

### 飞行器智能技术团队提出知识和数据融合驱动的队伍保持与协同避碰算法

发表日期: 2020-08-20 【大 中 小】 【打印】 【关闭】

## 1.什么是队形保持与协同避碰

多机器人编队队形保持与协同避碰任务，是指由多个单体机器人所组成的机器人编队，在作用空间下既要最大可能地保持所要求的队形，又要避免与环境障碍物及其他机器人发生碰撞，并最终到达目的地。该项任务对单体机器人的感知与控制能力和多机器人编队协同能力都有较高的要求，可广泛应用在诸多场景中，如智能仓储中的协作搬运、空域与海域的联合侦查探测等等。

经典的多机器人编队队形保持与协同避碰算法通常可以分为基于模式切换及引入势场力的方法。这些方法难以处理复杂动态任务，且存在目标不可达、行为输出局部震荡等弊端。近几年发展出的端到端学习方法（如多智能体强化学习方法）为该类问题的解决带来了新思路，但现有工作主要聚焦于避碰问题，在多机器人编队队形保持与协同避碰的融合问题中研究较少。

## 2.知识和数据融合驱动的队伍保持与协同避碰算法

中科院自动化所飞行器智能技术团队一直致力于无人自主系统的群体智能研究，近阶段将知识和数据协同驱动的群体智能实时推理与决策作为突破重点，融合规则知识、已有算法和模型等知识驱动方法以及深度强化学习、演化计算等数据驱动方法，实现群体的智能自主进化，并应用于虚实结合的集群验证系统。团队在最新的研究中，将机器人协同避碰与编队控制相结合，提出了一种基于模型知识和数据训练融合的多机器人编队队形保持与协同避碰算法，有效解决上述传统方法弊端和数学驱动训练时间长等问题，在性能表现方面取得了大幅提升。

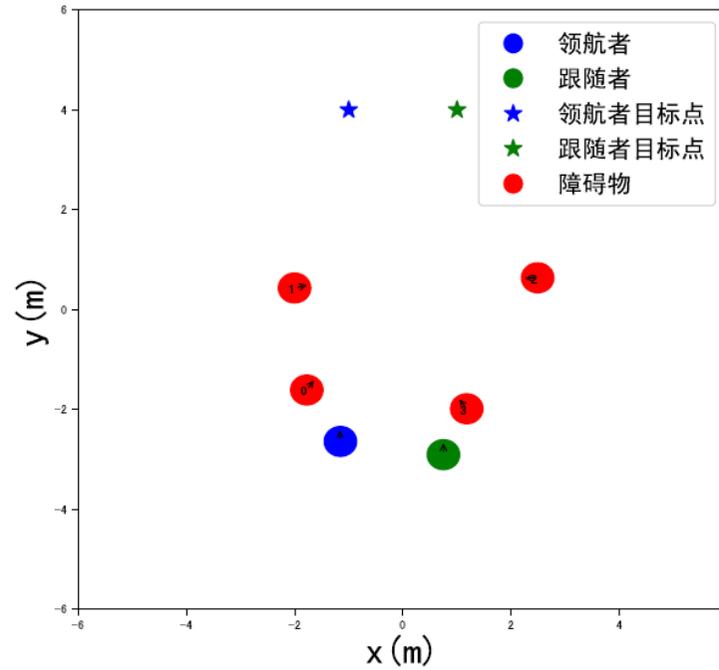


图1 问题示意

该工作以基于深度强化学习的机器人避碰算法为基础，将多机器人编队队形保持与协同避碰问题建模为具有复合奖惩的马尔科夫决策过程。在如何平衡编队保持与协同避碰问题上，没有人为设置限定条件，而是通过智能体与环境的交互，使之通过学习获得能够得到长期累计回报最大化的行为策略。

$$\begin{aligned} & \arg \min_{\pi(\tilde{s}_t^F, \tilde{s}_t^L, \tilde{s}_t^O)} \mathbb{E} \left[ t_g, \sum_0^{t_g} \| \mathbf{p}_t^F - \mathbf{p}_t^L - \mathbf{H}_t \| \right] (\mathbf{s}_0^F, \tilde{s}_0^L, \tilde{s}_0^O, \pi) \\ & s.t. \left| \mathbf{p}_t^F - \overline{\mathbf{p}}_t \right| > r + \bar{r}, \forall \mathbf{p} \in Others, \forall t \\ & \mathbf{p}_{t_g}^F = \mathbf{p}_g^F \\ & \mathbf{p}_t^F = \mathbf{p}_{t-\Delta t}^F + \Delta t \cdot \pi \left( \mathbf{s}_{t-\Delta t}^F, \tilde{s}_{t-\Delta t}^L, \tilde{s}_{t-\Delta t}^O \right) \end{aligned}$$

为了解决环境内障碍物数目不确定的网络泛化问题，算法对环境内的障碍物信息进行排序使之具备序列特性，而后引入LSTM模型对其进行统一处理，最终得到整体的状态表征。

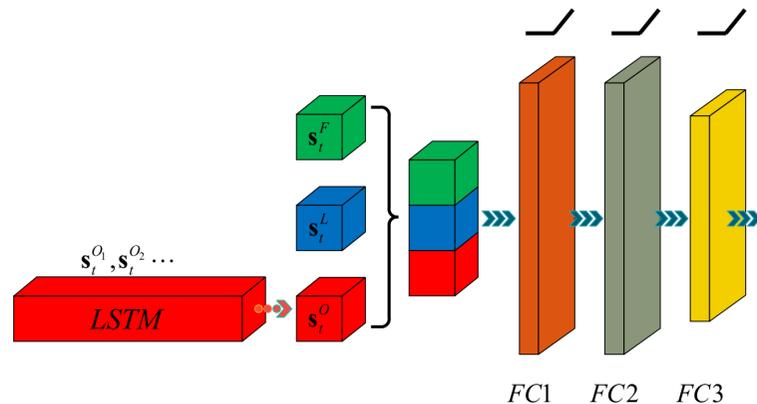


图2 基于LSTM的不确定数目障碍物状态表征

在模型知识与数据训练融合框架中，首先采用基于智能体运动学模型的知识驱动方法获取有效示例数据，基于该示例样本，反向计算环境状态与价值对，然后对其进行行为克隆。最终，在得到具有初始策略的网络后，将其参数赋予强化学习中的价值网络，并继续下一阶段的强化学习训练。

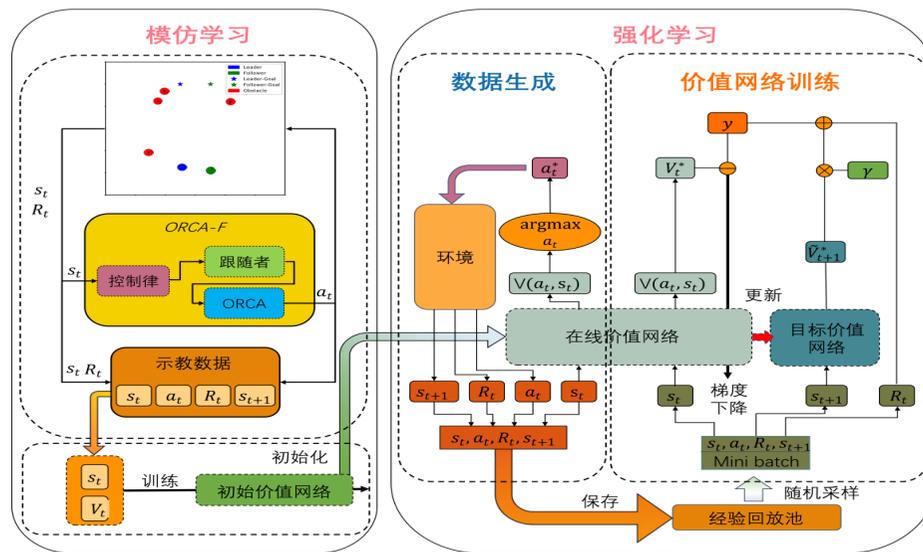


图3 算法总体框架

与传统的模仿学习-强化学习训练框架不同，该工作的模仿学习示例样本是采用基于智能体物理模型，且具有收敛性、最优性等理论保证的知识驱动方法生成的。其中，编队控制采用一致性理论设计分布式控制律，避碰策略则采用最优互补避碰（ORCA）算法来修正一致性编队控制器给出的智能体速度控制量。

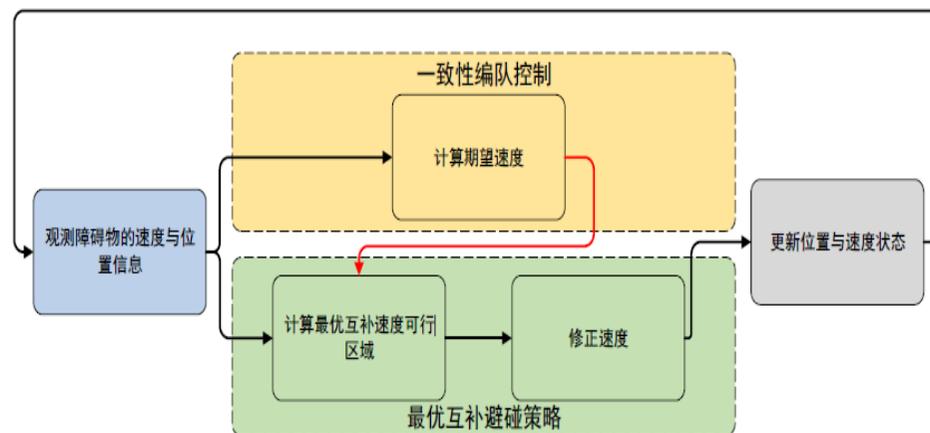
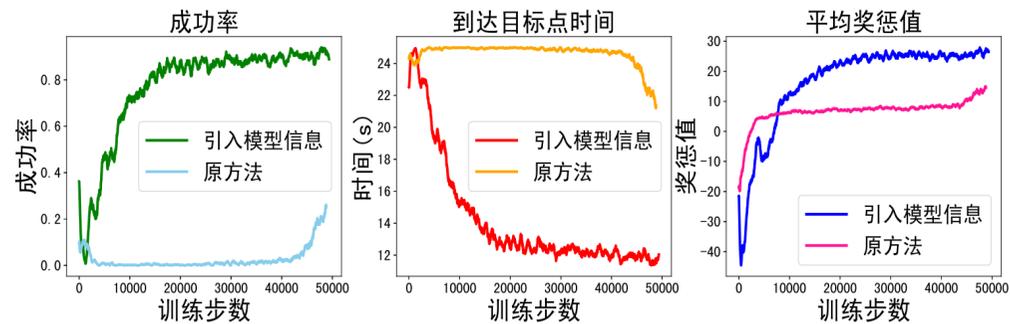


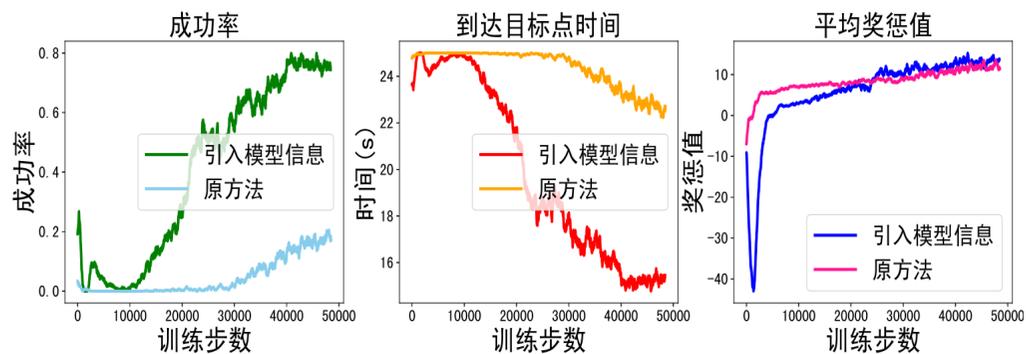
图4 知识驱动部分架构

### 3.对比与实测

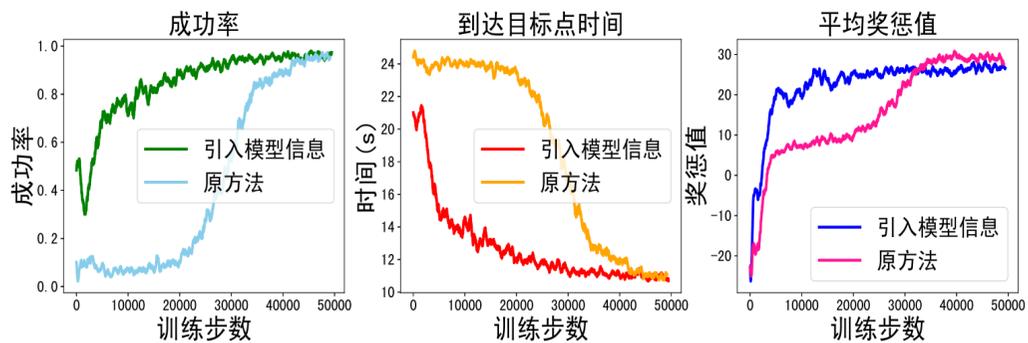
团队针对不同数目障碍物与不同目标队形，分别在四个场景下进行了仿真测试与验证。在验证基于知识驱动的模仿学习效果验证中可以看出，引入对模型驱动的编队控制与最优避碰方法的模仿学习后，多机器人编队无论是收敛速度还是最终结果，都取得了明显的提升。



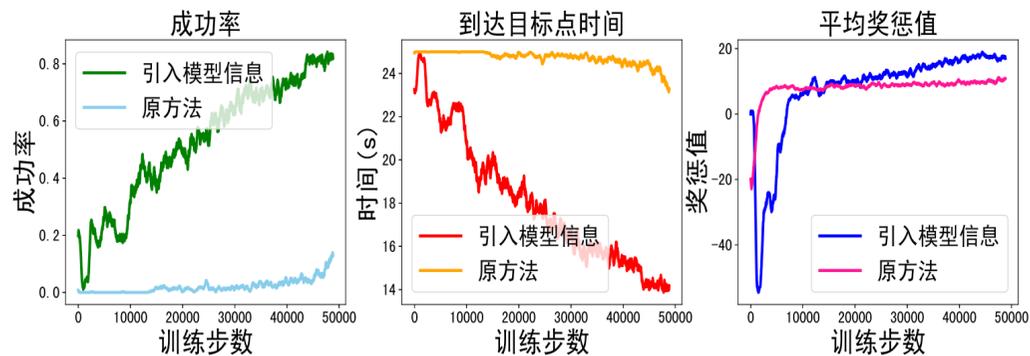
(a) 场景 1 训练曲线



(b) 场景 2 训练曲线



(c) 场景 3 训练曲线



(d) 场景 4 训练曲线

图5 仿真曲线

在此基础上，团队在每个场景分别进行了500组随机测试，并就其关键指标进行对比，发现知识和数据融合驱动的方法（FCCADRL）按编队队形无碰撞到达目的地成功率更高，并可在实际训练中极大缩减学习时间。

表1 算法对比效果

	Method	Nav time (s)	Success rate (%)	Average reward (-)	Danger frequency (%)	Average discomfort (m)
1	APF-F	11.31	72	20.4507	<b>10</b>	<b>0.15</b>
	ORCA	<b>9.47</b>	73	29.6348	62	0.09
	ORCA-F	9.7	83	<b>34.3366</b>	62	0.09
	FCCADRL	9.57	<b>92</b>	33.9282	17	0.14
2	APF-F	13.31	63	14.1197	12	<b>0.15</b>
	ORCA	<b>10.99</b>	57	15.8390	60	0.09
	ORCA-F	11.61	68	<b>22.4102</b>	67	0.09
	FCCADRL	11.33	<b>74</b>	17.4857	<b>5</b>	<b>0.15</b>
3	APF-F	11.04	68	25.5432	10	0.13
	ORCA	10.45	74	23.1645	23	0.09
	ORCA-F	10.30	75	25.6501	18	0.09
	FCCADRL	<b>9.67</b>	<b>98</b>	<b>31.5271</b>	<b>3</b>	<b>0.14</b>
4	APF-F	12.80	55	16.9860	11	<b>0.14</b>
	ORCA	12.11	62	12.3387	22	0.10
	ORCA-F	12.23	66	16.9721	18	0.10
	FCCADRL	<b>11.07</b>	<b>83</b>	<b>20.5463</b>	<b>5</b>	<b>0.14</b>

注：APF-F为引入势场力的方法，ORCA为最优互补避碰策略，ORCA-F为单纯采用本文所提出的示教学习的知识驱动方法，FCCADRL为本文知识和数据融合驱动的方法。

同时，团队在仿真基础上，搭建了一套基于UWB定位的室内多无人车集群系统，并在该系统中对所设计算法进行了实验验证。通过对照各场景下的轨迹图可以发现，无人车在柔性横队与纵队编队场景中均能成功完成任务，在保证运行安全的条件下尽可能地保持了编队队形，能够根据未知数目障碍物的未知运动主动调整动作，并最终到达目的地，验证了所设计算法的有效性。

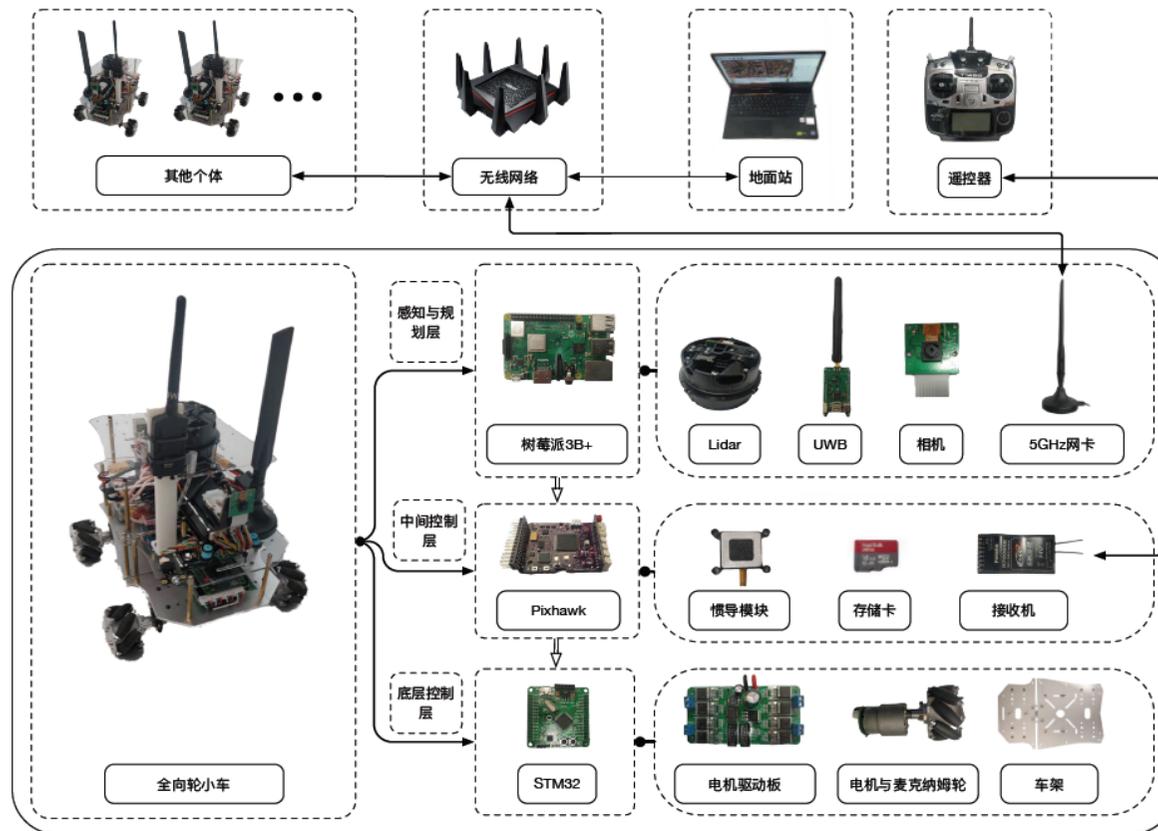
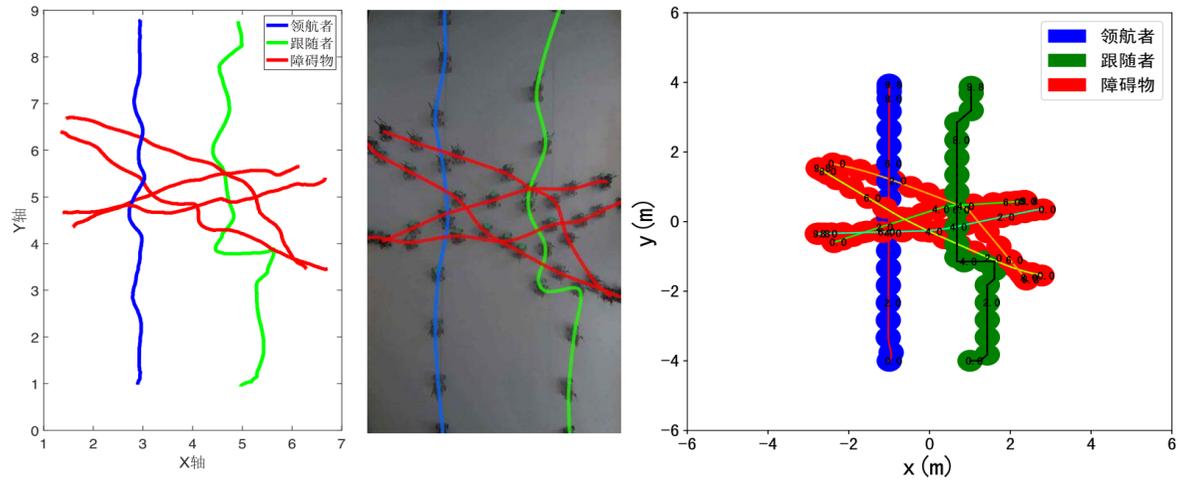
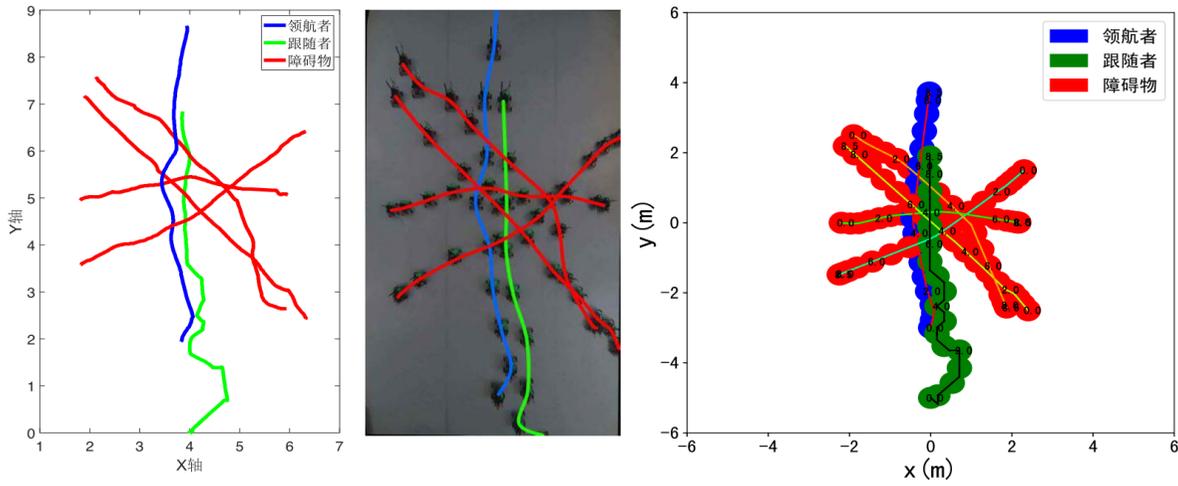


图6 试验平台结构组成



(a) 横队场景编队队形保持与协同避碰



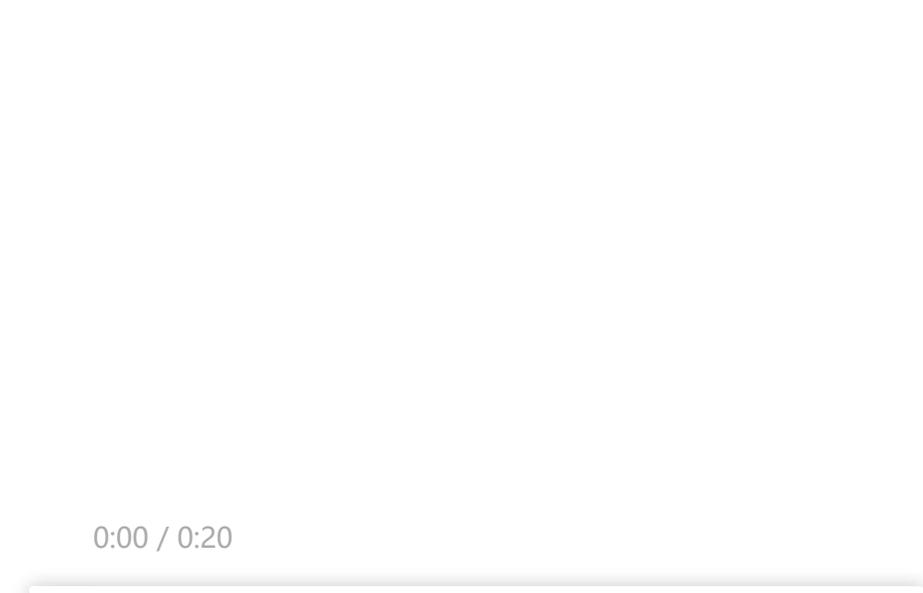
(b) 纵队场景编队队形保持与协同避碰

图7 实物验证

#### 4.总结

该研究在基于深度强化学习的避碰方法基础上，针对多机器人编队队形保持与协同避碰问题，提出了一种基于模型知识和数据训练融合的队形保持与协同避碰算法。针对无模型强化学习训练效率低、高度依赖数据，而实际机器人系统中往往存在一定的模型、经验知识可供利用的情况，通过对基于模型知识控制律的模仿学习，提升了深度强化学习的训练效率。该研究在仿真环境和实际环境对算法进行验证，相比于传统方法，取得了极大的性能提升。该方法有望解决实际物理集群系统面临的训练数据有限、存在物理及安全约束等条件下的高效学习问题，在仓储物流、无人码头配送等领域具有极大应用前景。

0:00 / 0:20



论文: Zezhi Sui, Zhiqiang Pu\*, Jianqiang Yi, and Shiguang Wu. Formation control with collision avoidance through deep reinforcement learning using model-guided demonstration. *IEEE Transactions on Neural Networks and Learning Systems*. 2020.

---



此网站支持IE9及以上浏览器访问

1996 - 2019 中国科学院 版权所有

备案序号: 京ICP备14019135号-3 (<https://beian.miit.gov.cn>) 京公网安备110108003079号

地址: 北京市海淀区中关村东路95号 邮编: 100190 Email: [casia@ia.ac.cn](mailto:casia@ia.ac.cn) (<mailto:casia@ia.ac.cn>)



([//bszs.conac.cn/sitename?method=show&id=08D8E9015DA3450AE053022819AC2F0E](http://bszs.conac.cn/sitename?method=show&id=08D8E9015DA3450AE053022819AC2F0E))