arXiv:2304.11204v1 [cs.RO] 21 Apr 2023

未知遮挡条件下分布式目标跟踪中的涌现协同行为

李天祺[1], 卢卡斯·W·克拉科夫[2] 和斯瓦米纳坦·戈帕尔斯瓦米[1]

摘要—利用多个机器人系统（MRS）跟踪多个移动关注对象（OOI）的问题，通常通过主动感知加以解决：主动感知维护关注对象的共享信念，并规划机器人的运动以最大化信息质量。机器人的机动性使其能够采取追求更佳可见性的行为，但这种行为会受到传感器视场（FOV）和遮挡物的限制。我们首先扩展了既有工作，以显式检测、维护和共享遮挡信息，从而即使缺乏 `先验语义遮挡信息，也能生成遮挡感知规划。主动感知方法的有效性通常依据估计误差和信息增益指标进行评估。然而，这些指标无法直接解释主动感知算法所促成的协同行为程度。随后，我们提取了源自相同底层算法、但在不同场景下呈现出不同形式的多种涌现协同行为。具体而言，我们重点阐述并展示了主动感知多机器人系统中的三种涌现行为模式：（i）当跟踪方向不同的轨迹时，或由于异构机器人之间的资源重新分配，智能体之间发生跟踪职责转移；（ii）感知智能体意识到轨迹受到遮挡，并在时间上暂时离开后返回；（iii）多机器人系统共享局部遮挡物信息，进而提升对遮挡的感知能力。

索引词—**主动感知、可见性、遮挡感知、涌现式协作行为、动态遮挡地图**

I. 引言

随着实际机器人应用的不断增长，感知的目的已从仅为机器人提供定位和导航所需的信息，发展为提供有关环境的更多上下文信息。利用机器人传感器跟踪环境中的不同移动目标具有重要意义。通常通过以下几个步骤完成目标跟踪：首先，原始传感器信号经过目标识别、目标定位和分类等处理层；其次，包含语义和位置信息的观测数据由数据关联算法进行处理，将观测结果与估计结果进行关联；第三，根据观测结果更新所维护的轨迹；最后，跟踪管理系统处理轨迹识别、删除、生成等情况。

当目标物体（OOI）不可观测时，使用上述方法进行跟踪会变得困难，因为

[1]Tianqi Li 和 Swaminathan Gopalswamy 任职于美国德克萨斯州大学城 77840 的德克萨斯农工大学机械工程系 {tianqili, sgopalswamy}@tamu.edu[2]Lucas W. Krakow 任职于美国德克萨斯州布赖恩 77807 的德克萨斯农工大学布什作战发展综合体 lwkrakow@tamu.edu 本研究得到美国陆军研究实验室的资助，合同编号为 W911NF1920243。本研究的维护地址为 https://github.com/ TianqiLi7398/emergent_behavior_active_tracking

当传感器的视场（FoV）受到限制，或者某些物体遮挡了感兴趣物体（OOI）时，便会出现这一问题。以往的研究主要通过在理解环境遮挡情况的同时，提高被跟踪物体的估计精度来解决这一问题。例如，视觉里程计通过多尺度全局注意力模块，在存在遮挡的情况下仍能对运动目标进行稳健的深度估计 [1]。在一种基于孪生网络的视觉跟踪方法中，局部上下文特征与全局语义信息相结合，以减轻遮挡和形变带来的干扰 [2]。关于被跟踪目标及其环境的额外语义信息，对于处理感知信号并将相关信息与被跟踪目标进行关联至关重要。针对人类运动的社会力模型，使联合概率数据关联滤波器（JPDAF）能够应用于存在智能体间交互和遮挡的人体跟踪场景。基于学习的变分贝叶斯聚类有助于估计包含被遮挡目标在内的目标数量 [3]。

上述方法仅关注目标跟踪的感知部分，而机器人本身的运动也可以进行规划，以提取与任务相关的信息（例如关注对象）。主动感知／主动目标跟踪 的研究受到智能动物行为的启发 [4]。主动感知机器人的一种实现形式是医疗保健地面机器人，它在封闭的关注区域（AOI）内跟踪人员 [5]。主动跟踪的规划器基于部分可观测马尔可夫决策过程（POMDP）构建。机器人不会贪心地将关注对象始终保持在其视场内，而是允许关注对象暂时离开视场，以平衡能耗与目标可见性。在另一个单机器人案例中，针对观察者—关注对象配对的追逃博弈展示了在时空坐标中进行最优轨迹规划的方法 [6]。

主动跟踪规划还可以扩展为考虑遮挡。可见性感知规划研究根据关注对象、遮挡物以及自体状态和配置来确定被检测的可能性。遮挡通常通过两种模式妨碍感知可见性：地面机器人视场中的水平不透明障碍物，例如墙壁 [7],[8], [6],人员 [5]和邻近车辆 [9]；无人机（UAV）垂直视场中树木 [10], [11], [12], 集装箱 [13],或其他机器人 [14]下方的阴影。上述方法依赖于关于遮挡位置的先验语义信息，因此能够改善状态估计。

多机器人系统可用于协同跟踪目标，以应对观测区域范围和目标数量不断增加的情况。将主动感知从单机器人平台扩展到多机器人系统（MRS），

催生了多项有趣的研究，例如对主动协同感知 [15]，信息驱动控制 [16],及其后续应用的研究，包括机器人救援 [17],事件记录与分析 [18], [19], 监视与探索 [20], 等。MRS 算法可以像单机器人情形一样扩展，以考虑遮挡。文献 [12]分析了树木遮挡的大小和密度对分布式目标跟踪（DTT）的影响。

多机器人系统并非多个单机器人系统的简单集合。随着智能体在作业区域内移动，其视场之间的冗余为协作和提升跟踪性能提供了重要机会。已有研究提出了明确的协作算法，例如在存在遮挡的复杂环境中，为最大化可视性而采用基于区域的方法（将关注区域划分为机器人的视场）和目标跟随行为模式 [8]。此外，文献 [21] 还提出了队形控制的思路，例如由具备冗余感知能力的机器人包围目标。

另一方面，以共同最小化目标为重点的隐式算法也已证明取得成功——相关实验 [10], 和仿真 [12] 展示了无人机团队在各种树木遮挡场景中的表现。POMDP 和模型预测控制（MPC）等通用前瞻方法已成功应用于主动感知 [5], [9], [11], [12], [22]。虽然目标函数能够衡量算法的效能，但识别协同行为并非易事。行为层面的分析很有价值，因为这使得人类推理能够与算法决策协同发挥作用。以往关于机器人和车辆的行为提取与分析工作包括：从行为视角定义不同的反应式控制器，见 Braitenberg 车辆模型 [23],；对牧群、鱼群和鸟群中的行为进行分类 [24]。机器人足球比赛中，通过行为仲裁触发诸如传球之类的协同行为 [25]。在多智能体牧羊任务中，观察到双方围绕绵羊同步振荡的耦合振荡围控行为更为有效 [26]。以往对机器人行为的分析促使我们在主动目标跟踪中定义涌现协同行为，而这一主题此前尚未得到研究。

本文的主要贡献如下：

- 我们引入了动态遮挡地图 的概念，用于维护静止遮挡对象（SOO）。无人机同时检测、跟踪并共享关注对象（OOIs）和 SOO 的信息，并针对它们采用不同的运动假设。在缺乏遮挡先验语义信息的情况下，动态遮挡地图可通过提高机器人运动规划对“刚检测到的”遮挡的感知能力，降低 DTT 的跟踪性能退化。

- 协同行为的明确定义：我们将其作为主动目标跟踪所产生的涌现行为加以展示。

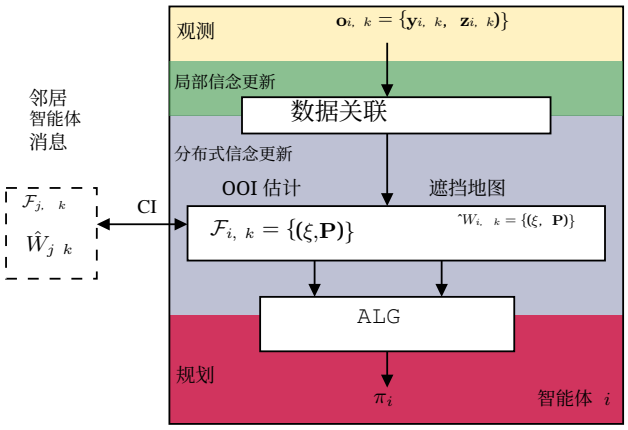


图 1：具有动态遮挡地图 W 的 DTT 范式中的遮挡感知主动感知。

II. 未知遮挡下多传感器多目标跟踪中的主动感知

在本文中，我们使用字母 i, j 作为智能体的索引；h, k 作为离散时间索引； N 作为团队中智能体的总数；整数集合 $[N] = \{1, 2, ..., N\}$；在多机器人系统中，智能体和机器人这两个词可互换使用，均表示决策单元。协同 N-机器人目标跟踪被纳入POMDP框架，可表示为元组 $\mathcal{P} = (\mathcal{S}, \mathcal{A}, \mathcal{T}, \mathcal{O}, C, \Omega\ \gamma)$，与 [5], [12], [27]类似。此外，决策过程中通常会考虑规划时域 H 。目标跟踪场景如下：总计 N 架无人机以固定高度飞行，并在一个平面的二维封闭关注区域内跟踪关注对象；每架无人机均配备一台镜头垂直向下、朝向关注区域地面的相机。多机器人系统拥有一组关注对象和SOO的语义标签，例如 $\mathcal{L}_{\text{OOI}} = \{\text{人}\}$, $\mathcal{L}_{\text{SOO}} = \{\text{树}\}$。我们假设环境中包含一组 n SOO $W = \{w_1, ..., w_n\}$，其中每个元素都存在于关注区域的封闭范围内。

我们假设团队中任意一对无人机之间都具有理想的信息共享连通性。每个智能体都运行带线性卡尔曼滤波器的 JPDAF 来处理本地观测，并通过信息共识（CI）算法融合信息，以维护 [11]中描述的全局 DTT。

状态空间 $\mathcal{X} = (\mathcal{S} \times \chi \times \mathcal{F})$ 包含：联合机器人状态 $s \in \mathcal{S}$，其中 $\mathcal{S} = \Pi_{i=1}^{N} \mathcal{S}_i$；以及第 i 个单独机器人状态 $s_i \in \mathcal{S}_i$，其定义为 $s_i = (p_i^x, p_i^y, \psi_i, v_i^x, v_i^y) \in \mathbb{R}^5$，位于二维水平面内，包括位置 p、偏航角 ψ 和速度 v；传感器的视场 $\phi_i(s)$ 是由机器人状态和传感器规格参数化的投影感知区域，因此团队的传感器覆盖范围由其视场构成的元组描述， $\phi(s) = (\phi_1, ..., \phi_N)$；关注对象的状态 χ 包含所有关注对象的位置和速度；滤波器（估计）状态 $\mathcal{F}$ 是由基于共识的 DTT 算法 [11] 维护的高斯分布 $\mathcal{N}(\xi, \mathbf{P})$。

联合**动作**空间 $\mathcal{A} = \Pi_{i=1}^{N} \mathcal{A}_i$ 描述发送给团队的水平速度指令，其中每个机器人

$a_i = (a_i^x,\ a_i^y) \in \mathbb{R}^2$ 和速度约束 $|a_i| \le v_{max,\ i}$。

观测中的不确定性**观测** 带来了观测空间 Ω 和 POMDP 中的观测规则 $\mathcal{O}$：传感器获得观测 $\mathbf{o} = (\mathbf{y},\mathbf{z}) \in \Omega$，其中包含对象的语义标签 $\mathbf{y}$ 以及传感器视场中的对象$\mathbf{z}$ 的位置。传感器模型假设不确定性仅存在于位置部分 $\mathbf{z}$中。对于一对对象—观测者 (t,i)，我们用字母 t 表示 OOI 的索引，用 i 表示机器人的索引。观测定律 $\mathcal{O}(x|o)$ 是 OOI 状态 χ下的二重条件概率。

1) **定位** 在时间步 k中的噪声：

$$\mathbf{z}_{t,i,k} = \mathbf{H}_k\chi_{t,k} + \mathbf{v}_{t,i,k}, \mathbf{v}_{t,i,k} \sim \mathcal{N}(0, \mathbf{R}_{t,i,k}) \quad (1a)$$

$$\mathbf{R}_{t,i,k} = \alpha_i \mathbf{G}(\rho_{t,i,k}) \begin{bmatrix} 0.1d_{t,i,k} & 0 \\ 0 & 0.1\pi d_{t,i,k} \end{bmatrix} \mathbf{G}(\rho_{t,i,k})^T \quad (1b)$$

在 (1a) 中，由机器人 i 检测到的 t 位置包含高斯白噪声 $\mathbf{v}_{t,i,k}$。协方差$\mathbf{R}_{t,i,k}$ 是由传感器与 (1b) 中 OOI 之间的欧氏距离 $d_{t,i,k}$ 和方位角差 $\rho_{t,i,k}$所确定的距离—方位 感知噪声，以及标量 α_i。

2) **可见性**：当且仅当对象 t 在时间 k 处于机器人的 i 视野内且不在任何遮挡区域中时，该对象可见；机器人 i 可见对象的集合定义为

$$\mathcal{V}_{i,k} = \{t|(\chi_{t,k} \in \phi_{i,k}) \wedge (\chi_{t,k} \notin W)\}, \quad (2)$$

以及团队可见的物体为 $\mathcal{V}_k = \cup_{i\in[N]} V_{i,\ k}$。

信念状态空间 $\mathcal{B}(\mathcal{X})$ 在目标跟踪问题 [12] 中，通过概率分布描述部分可观测的 OOI 状态 ξ。信念状态 $b_k(x) := b(x_k) \in \mathcal{B}(\mathcal{X})$ 包含状态 $b_k(x)$ 中每个组成部分的信念，即 $= P_{x_k}(x|Z_0, Z_1, ..., Z_k, u_0, ..., u_k) = (b_k^s, b_k^\chi, b_k^\xi, b_k^P)$：可观测状态包括机器人状态 s 和 OOI 轨迹的滤波状态 $\mathcal{F}(b_k^s, b_k^\xi, b_k^P)$，可由狄拉克 δ 函数表示，例如 $b_k^s = \delta(s - s_k)$；目标信念状态是由滤波状态 F_k 表示的条件概率分布，这仅适用于此特定实现。

信念状态**转移** $b_{k+1} \sim \mathcal{T}(\cdot|b_k, a_k)$由与三个信念状态元素相对应的分量构成：i) 从机器人状态 b_k^s到机器人的确定性速度指令 a_k；ii) OOI状态 χ_k 具有随机性，且独立于传感器状态 s_k和控制变量 a_k；iii) 滤波器轨迹估计中的近似恒定速度（*NCV*）运动模型：对于一个状态为$\xi_{t,k} = (p_{t,k}^x, p_{t\ ,k}^y, v_{t,k}^x, v_{t\ k,}^y) \in \mathbb{R}^4$的OOI，其动力学假设为

$$\xi_{t,k+1} = \mathbf{F}_k\xi_{t,k} + \mathbf{w}_k, \mathbf{w}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}). \quad (3)$$

我们使用线性卡尔曼滤波器和 NCV 运动模型，根据 $\mathcal{L}_{\text{OOI}}$ 中的语义更新估计轨迹。

我们考虑代价函数 $C: \mathcal{B}(\mathcal{X}) \times \mathcal{A} \to \mathbb{R}_{\ge 0}$，它表示关注对象的总期望均方跟踪误差。该代价函数等于在给定信念状态转移的情况下，后验协方差矩阵 $\mathbf{P}$ 的迹在高斯噪声假设下 [22],

$$\begin{aligned} C(b_k, a) &= \int \mathbb{E}_{w_k, v_{k+1}}\left[\ ||\chi_{k+1} - \xi_{k+1}||^2 | x_k, a\right] b_k^\chi(x)dx \\ &= tr[\mathbf{P}_{k+1|k+1}] \end{aligned} \quad (4)$$

一个**策略** $\pi: \mathcal{B}(\mathcal{X}) \to \mathcal{A}$ 根据一个信念状态生成动作选择。在这种协同跟踪情况下，**目标**在 H步前瞻滚动规划时域内累积代价 $r_k = C(b_k,\ a_k)$。采用策略 π 时信念状态 b的价值，是滚动规划时域 H 内的期望累积代价，如下式所示。

$$J_H^\pi(b) = \mathbb{E}_{\mathcal{T}(\cdot|b_k, a_k)}\left[\sum_{k=0}^{H-1} \gamma^k r_k \middle| a_k = \pi(b_k), b_0 = b\right] \quad (5)$$

通过展开每个智能体的策略，目标函数可以等价地表示为$J_H^\pi(b) = J_H(b, \pi_1, ..., \pi_N)$。

III. 未知遮挡下的 DTT 规划

我们将生成目标（5）所定义策略的一般规划算法记为 ALG。直接的方法是采用集中式 ALG 来直接求解优化问题。纯分散式规划 Dec-POMDP 消除了对其他智能体通信的所有依赖，因此每个智能体 i 都会根据局部轨迹估计 $\mathcal{F}_i$ 生成最优计划或宏动作 [28]，其中

$$\pi_i^{\text{Dec-POMDP}} \in \arg\max_\pi J_H^\pi(b; \mathcal{F}_i) \quad (6)$$

在 DTT 中应用共识算法，可保证智能体之间的差异有界 $\mathcal{F}_i$ [11],，从而使分散式部分可观测马尔可夫决策过程的解与集中式 POMDP 的解相同。

团队规模的增加 N 会使 (6) 的计算复杂度呈指数级增长。从实现角度来看，更易处理的 ALG 更受青睐，尽管这会牺牲最优性。这里，我们采用一种基于逐个智能体规划的 SMA-NBO 方法，该方法基于意图 [12]。将智能体 i 的意图记为 $\bar{\pi}_i$。SMA-NBO 在每个时间步生成一个计划，$k = 1$为 $\left(a_1,\ a_2,\ ..a_H | a_k = \pi(b_k)\right)$。初始动作 a_1 立即作为控制量执行，这与其他滚动时域方法一致。然而，在下一个决策阶段 $k = 2$，SMA-NBO 会利用剩余动作序列，并将其作为未来意图传达给其他机群成员，以便为下一个决策时刻提供信息，即 $\bar{\pi} := (a_2,\ ..a_H)$。策略 π_i^{SMA}是基于 1 的单智能体策略改进：$i-1$ 智能体更新后的策略$\pi_1^{\text{SMA}}: i-1$ 以及其余智能体的意图 $\bar{\pi}_{i+1:N}$。

$$\pi_1^{\text{SMA}} = \arg\max_{\pi_1} J_H(b, \pi_1, \bar{\pi}_{2:N}) \quad (7a)$$

$$\pi_i^{\text{SMA}} = \arg\max_{\pi_i} J_H(b, \pi_{1:i-1}^{\text{SMA}}, \pi_i, \bar{\pi}_{i+1:N}) \quad (7b)$$

$$\pi_N^{\text{SMA}} = \arg\max_{\pi_N} J_H(b, \pi_{1:N-1}^{\text{SMA}}, \pi_N) \quad (7c)$$

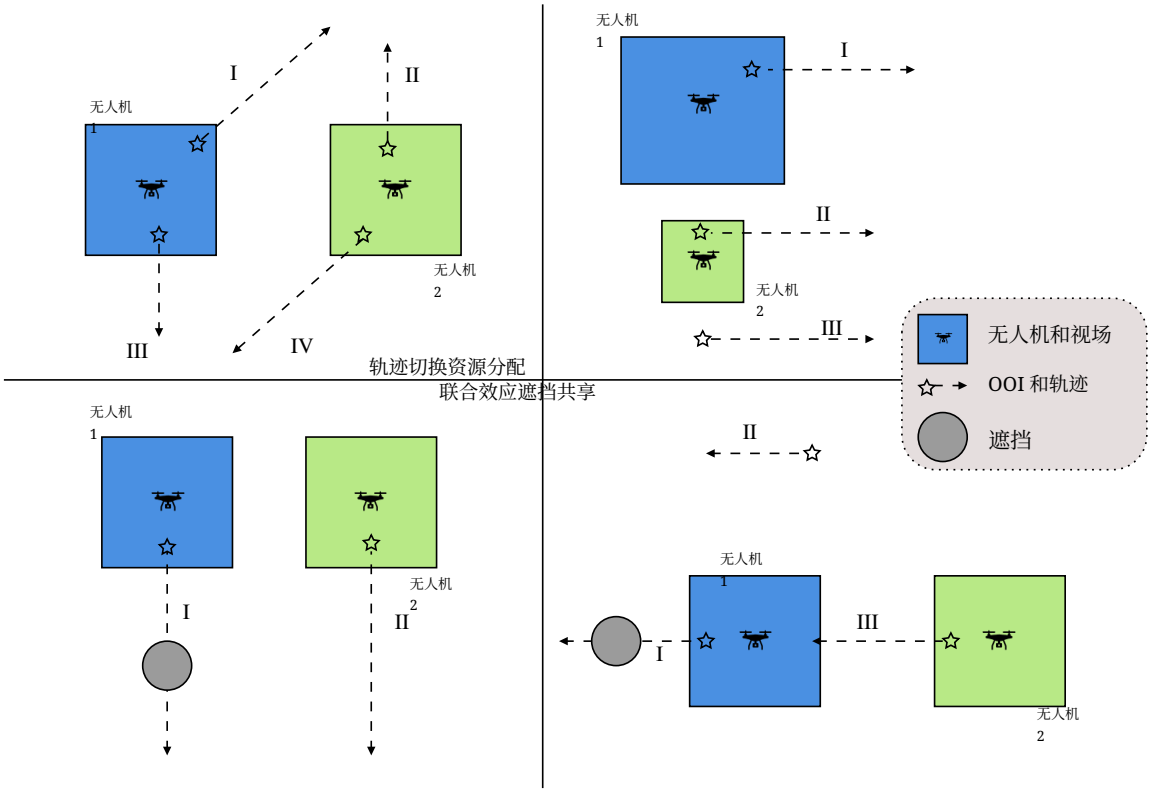


图2：具有潜在协同行为的 DTT 案例。在这四种情况下，轨迹真实值以虚线箭头表示。每种场景中标记的所有关注对象均以相同的不确定性矩阵$\mathbf{P}$ 初始化进行轨迹估计 $\mathcal{F}_0$。每种场景的持续时间，L，各不相同。

A. 未知遮挡

如第二节所述，遮挡感知规划要求规划器在前瞻时域内检查所有 OOI-观察者对 $(t,\ i)$ 的可见性。在这个特定问题中，有效的 `先验遮挡地图 W 有助于 J_H^π 信息滤波器格式 (4) 中协方差矩阵的更新：

$$\mathbf{P}_{k+1|k+1,t} = \left(\mathbf{P}_{k+1|k,t} + \sum_{t\in\mathcal{V}_k} \mathbf{H}_k^T \mathbf{R}_{t,i,k}^{-1} \mathbf{H}_k\right)^{-1} \quad (8)$$

如果 `先验地图 W 无效或不可用，ALG 的行为将不再具备遮挡感知能力。

然而，如果未知的 SOO 能够通过观测被识别，就可以在图1所示的范式中通过 DTT 加以维护。这里我们假设，对于 OOI 和 SOO，观测定律在传感噪声模型 (1b) 和可见性 (2) 上是一致的。在每个时间步，数据关联算法 JPDAF 会更新标签和位置的局部观测元组 $(\mathbf{y}_k,\mathbf{z}_k)$。可以对 SOO 应用语义感知运动模型 (3)，例如，将*tree* 物体假设为静态，即$\mathbf{F}_k = \mathbf{I},\mathbf{Q} = \mathbf{0}$。SOO 的估计值 ˆ 构成动态遮挡地图$W_{i,k}$，并由 CI 共享和更新。DTT 框架使所有 ˆ 多机器人系统成员都能获取关于全局 SOO 的时序知识。更新后的地图 $W_{i,k}$ 会立即用于 (2)，以便在规划 ALG 中近似 W 。尽管动态遮挡地图仅包含已检测到的 SOO 估计值，但如果某个 OOI.....它会提升对近期时间步的遮挡感知能力。

正在朝着 SOO 估计迈进。此外，共享 SOO 估计可增强多机器人系统中的协作，相关分析见第 IV-C 节。

IV. 主动感知中的涌现行为分析

本节定义 DTT 任务中的涌现式协同行为，并对图2中的 DTT 场景进行概括。这些场景包括两种无遮挡情形（轨迹交接和资源分配），以及两种存在遮挡的情形（联合效应和遮挡共享），以阐明主动目标跟踪中的涌现式协作。所有场景均按照以下设置进行仿真。

A. 仿真设置

仿真的主要目的是为涌现行为分析提供实证统计数据。所有数值仿真均在德克萨斯农工大学高性能研究计算中心提供的先进计算资源上进行 [1]。

规划算法 SMA-NBO 和采用 NBO 的分散式部分可观测马尔可夫决策过程均已实现，以便比较和对照它们的合作倾向，尽管二者的共同目标函数并未明确规定任务协作要求。两种算法都使用名义信念轨迹。

[1]TAMU HPRC：https://hprc.tamu.edu/

在轨迹信念状态（4）上的代价函数近似方法，该方法通过不含噪声 [12]的运动模型（3）预测 OOI轨迹。该名义信念轨迹被视为 NCV模型中高斯假设下的最大似然轨迹。粒子群优化（PSO）在两个算法中均有实现，用于对目标函数 4 进行数值优化。

在轨迹交接和资源分配中，我们将 ALG 的前瞻时域设为 $H=1$ 和 5。 $H=1$规划时域提供了行为分析中的基准贪心解。两种情景均在 ALG 与 H 的每种组合下重复进行 40 次试验。

在图1所示的主动感知范式中，遮挡地图 W 是触发遮挡相关行为的关键组成部分。联合效应和遮挡共享情景展示了这些行为。ALG 固定为 SMA-NBO 和 $H=5$，并针对不同的遮挡地图设置重复进行 40 次仿真： `先验遮挡地图、动态遮挡地图和无遮挡地图。我们从行为视角研究动态遮挡地图的有效性。

B. DTT 中的协同行为

多机器人系统的冗余性和异质性使主动感知中的协同行为成为可能。与探索导向任务不同，目标跟踪目标会促使机器人采取以感兴趣目标为中心的行动，以增加目标的可见机会，从而降低感兴趣目标估计的定位不确定性。多个方向各异的感兴趣目标轨迹，使单机器人情形下这种“持续观测”行为面临挑战。有限的视场要求在各条轨迹的可见性之间取得平衡，这一点可从式（4）中看出。在 DTT 中，机器人通过将团队成员职责分配给不同方向上对感兴趣目标估计的任务来克服这一挑战。为识别这种职责，我们定义*m*-所有权，以分析多机器人系统在整个仿真时域内的行为。

定义IV.1 (*m*-归属(i,t,k))。 机器人 i 在时间步 k 获得目标 t 的*m*-归属，当且仅当机器人 i 连续 m 个时间步检测到 t，即 $\forall 0 \le h \le m,\ t \in \mathcal{V}_{i,\ k+h}$。

令 $M_k=(M_{1,k},\ldots,M_{N,k})$ 表示时间索引为 k 时所有 N 个机器人的*m*-归属，其中每个子集 $M_{i,k}$都包含机器人 i拥有的 m 个目标索引集合。然后，我们定义所有权重新分配行为。

定义IV.2 (所有权变更$[k,i \xrightarrow{t,L} j]$)。 机器人 i之间会发生所有权变更 行为: j: 在时间步 0，目标 $t \in M_{ik} \wedge t \notin M_{jk}$，经过 L 个时间步后，所有权转移至 j，即 $t \in M_{jk+L} \wedge t \notin M_{ik+L}$。

所有权变更 描述了一对机器人 i之间目标t 的可见性变化， j。图2中的轨迹交接场景就是一个例子。在时间 $k=0$， *m*- 所有权 $M_0=\{M_{\text{UAV1},\ 0}$：(I, III)$M_{\text{UAV2},\ 0}$：(II, IV)$\}$。轨迹估计 $\mathcal{F}_0$ 预测了两架无人机（I 和 II、III 和 IV ）所拥有轨迹的分歧。当相同无人机所拥有的轨迹经过 L 个时间步后.....

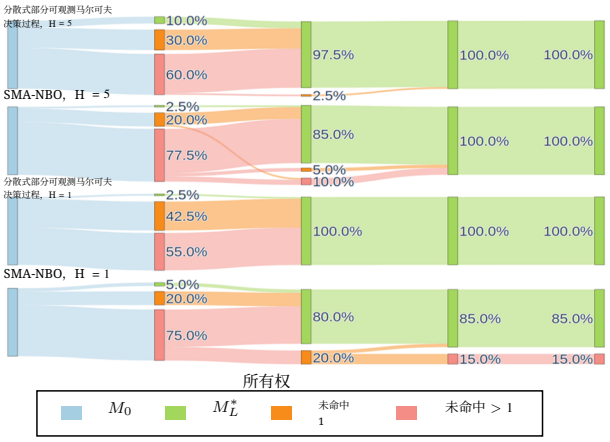


图 3：轨迹交接场景中 OOI 所有权的桑基图。

在距离上充分分离时，直觉上认为，最优的合作结果应采用“区域防守”策略。更简洁地说，期望得到的所有权分配为 $M_L^*=$ ((I, II)，(III, IV))，其中未指定机器人，因为机器人索引的两种排列所代表的合作水平是等价的。从 M_0 到 M_L^*，所有权变更 会发生在以下两种情况之一中。

$$(0,\text{UAV1} \xrightarrow{\text{I},L} \text{UAV2}) \wedge (0,\text{UAV2} \xrightarrow{\text{IV},L} \text{UAV1})$$
$$(0,\text{UAV1} \xrightarrow{\text{III},L} \text{UAV2}) \wedge (0,\text{UAV2} \xrightarrow{\text{II},L} \text{UAV1})$$

轨迹交接的仿真分析：图 3 绘制了整个仿真过程中*5-* 所有权 的变化情况。在每个桑基图中，均可视化整个仿真过程中*5-*所有权 状态的变化：从左到右，每一列表示在多次重复试验中，时间步 k =0, $\frac{L}{4}$, $\frac{L}{2}$, $\frac{3L}{4}$, L 时*5-*所有权状态出现的频率。图例“未命中 1”和“未命中 $>$ 1”分别表示任意无人机不拥有任何关注对象或拥有超过 1 个关注对象的所有权状态。在轨迹交接过程中，无人机难以在 $k=\frac{L}{4}$ 时刻维持原有的*5-*所有权 。从 $k=0$ 到 $k=\frac{L}{2}$ 或 $\frac{3L}{4}$ 可以观察到所有权变更 行为，此时无人机会暂时离开部分位于视场之外的关注对象轨迹（ $k=\frac{L}{4}$ 和 $\frac{L}{2}$ 时刻的橙色和红色节点）。比较不同算法的桑基图可以发现，分散式部分可观测马尔可夫决策过程总体上具有更强的协作性。在 $H=5$ 情况下，分散式部分可观测马尔可夫决策过程和 SMA-NBO 均实现了 100% 的 M_L^*，但分散式部分可观测马尔可夫决策过程在早期阶段执行 M_L^* 的频率更高，例如，在仿真进行到一半时，两者的比例分别为 97.5% 和 85%。随着规划时域的增加，SMA-NBO 的改进更加显著，因为短视的 SMA-NBO（ $H=1$）有 15% 的仿真试验最终停留在协作性较低的*5-*所有权状态。对所有权变更 行为的分析证实，与分散式部分可观测马尔可夫决策过程相比，计算效率更高的 NBO 的 SMA 实例化方法在最优性方面略有下降。此外，贪心行为（$H=1$）与非短视行为（H =5）之间的对比表明，在轨迹交接过程中，通过前瞻规划隐式获得了协作收益。

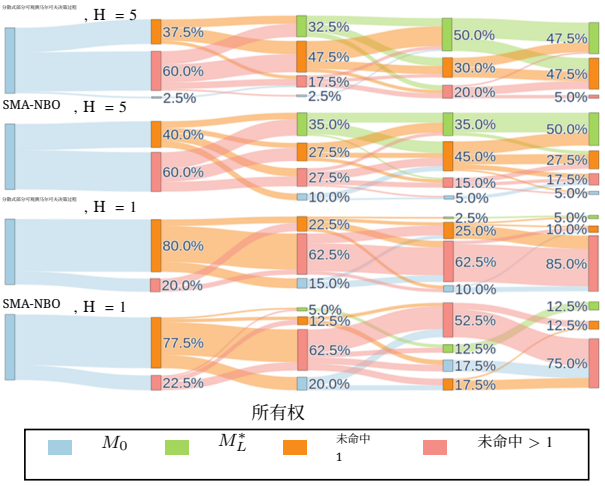


图4：资源分配场景中的OOI所有权桑基图。

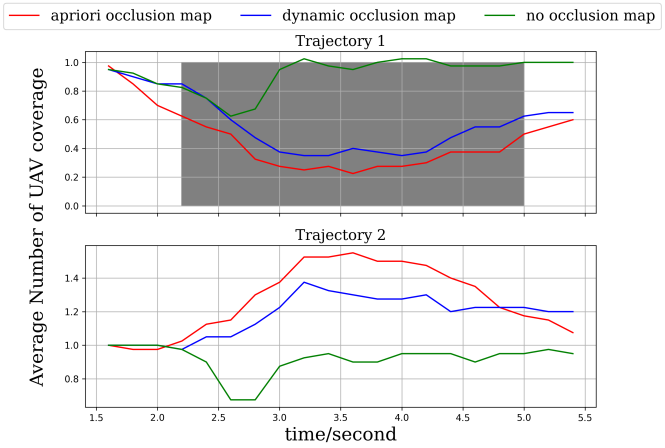


图5：联合效应场景下的轨迹覆盖。灰色区域表示该轨迹处于遮挡区域内。

场景。

除了跟踪发散轨迹外，协同所有权变更 也可能因多机器人系统的异质性而发生。图2中的资源分配场景包含视场大小不同的无人机，且两架无人机都无法看到目标III。如果保持相同的所有权，无人机2需要以之字形模式移动，以跟踪轨迹II和III，这一点也在 [22]的图1中有所报告。通过交换位置，最优的协同行为是$M_L^* = (M_{\text{UAV1},\ L}:$ (I) $M_{\text{UAV2},\ L}:$ (II、III)$)$。将视场资源分配给 M_L^* 是协作性的，因为关注对象的可见性得到了改善。

资源分配的仿真分析：与图 3 类似，图 4 展示了重复仿真的*3*-所有权结果。在此场景中，所有权变更行为比轨迹交接更具挑战性；只有大约一半的仿真在规划时域为$H = 5$时实现了M_L^* 。在相同规划时域下，Dec-POMDP 和 SMA-NBO 的*3*-所有权流达到合作状态的概率相近。对于贪心规划时域 $H = 1$，底部两个桑基图中的规划器分配资源的概率都很低（Dec-POMDP 中为 5%，SMA-NBO 中为 12.5%）。与轨迹交接中从 M_0 到 M_L^* 每架无人机都拥有一个未改变的目标不同，在资源分配中，两架无人机从M_0 到 M_L^* 的*3*-所有权完全不同。要实现合作式所有权变更，需要具备足够的前瞻规划时域并避开所有局部最优，这对于数值方法而言是一项严苛的任务。因此，该场景中的合作表现并不明显。不过，更长的前瞻规划时域展现出合作倾向。

C. 存在遮挡的协同 *DTT*

对遮挡的感知为 DTT 任务带来了独特的涌现式协同行为。一种通用的遮挡感知推理方式是：如果某条轨迹 t 将在未来时间区间 $[h, L-h]$内被遮挡，则应在此区间内将感知资源分配给其他轨迹。

定义 IV.3 (遮挡感知(k_0, t, W, L, h))。 关注对象t 在所有时间步 $k \in [k_0 + h, k_0 + L - h]$, $\chi_{t,k} \in W$均被遮挡。满足以下三个条件时，我们称该行为具有遮挡感知 能力。

1) $\exists S_1 \subseteq [N]$, $S_1 \neq \emptyset$ s.t. $\forall i \in S_1$, $t \in M_{k_0,i}$;
2) $\nexists i \in [N]$ s.t. $\forall k \in [k_0 + h, k_0 + L - h]$, $(M_{k,i} = \emptyset) \wedge (\chi_{t,k} \in \phi_{i,k})$;
3) $\exists S_2 \subseteq [N]$, $S_2 \neq \emptyset$ s.t. $\forall i \in S_2$, $t \in M_{k_0+L,i}$;

在遮挡感知 的定义中，第一个和第三个条件确保关注对象 t 在遮挡时段 $[k_0+h, k_0+L-h]$之前和之后至少对一个机器人可见；第二个条件指出，当 t 被遮挡时，所有机器人都应寻找其他关注对象，而不是在t 是其视场内唯一 的关注对象时仍绝望地等待 t ，这体现了对遮挡的感知。

一个例子是图2中的联合效应场景。轨迹I在时间区间$[h, L-h]$内进入遮挡区域，且初始位置的*m*-归属 为 $M_0 = (M_{\text{UAV1},0}$: (I)$M_{\text{UAV2},0}$: (II)$)$。此处遮挡感知 协同行为要求两架无人机共同覆盖轨迹II，直至轨迹I离开遮挡区域，即 $\forall k \in [h, L-h], M_k = (M_{\text{UAV1},k}:$ (II) $M_{\text{UAV2},k}:$ (II)$)$。在最后阶段，一架无人机覆盖轨迹I，$M_L = ($(I),(II)$)$。

联合效应的仿真分析：图5展示了联合效应的仿真结果。图中按轨迹绘出了仿真时域内无人机覆盖（*1*-归属）的平均数量。2.2 ∼ 5.0 秒之间的灰色区域表示轨迹 I 处于遮挡状态的时间段。从图中，我们比较了 SMA-NBO 在掌握不同遮挡地图信息时的行为：`先验、动态和无遮挡地图。无遮挡地图下的行为有明显差异，因为在遮挡期间，它仍持续尝试观测轨迹 I。`先验地图和动态地图显著降低了轨迹 I 的

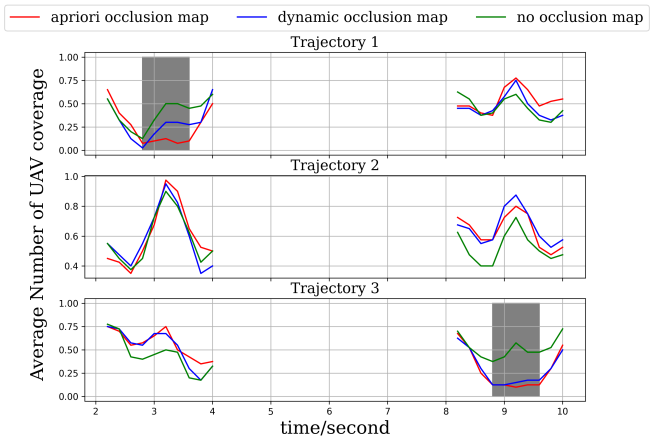


图 6：遮挡共享场景下的轨迹覆盖。灰色区域表示该轨迹处于遮挡区域内。

在灰色区域内收敛，同时有多架无人机覆盖轨迹 II。实证结果表明，ALG 在使用 `先验和动态遮挡地图时表现出遮挡感知行为。然而，与红线相比，蓝线在 1.6 ∼ 2. 2 秒处出现了明显的变化延迟。动态地图需要先检测到遮挡，才能生成遮挡感知规划，这解释了蓝线响应较慢的原因。

动态遮挡地图基于环境探索运行，其中包含一种特定形式的协同行为。一旦机器人 i 检测到新的 SOO，就会将此局部信息分享给多机器人系统中的所有机器人，如图1所示。这种知识共享是一种协同行为，可提升团队对*SSO* 的认知。

定义 IV.4 (遮挡共享)。 定义一次遮挡检测事件 e 为在时间步 k，机器人 i 检测到遮挡 $w \in W$，即 $e = \{w \in \phi_{i,k}\}$，则 $\forall h > k$ 以及任意轨迹 t。

遮挡共享 的定义大致基于未来发生行为事件遮挡感知 的概率，

$$\mathbb{P}[OA(t, W_h, L, h)|e] \geq \mathbb{P}[OA(t, W_h, L, h)], \qquad (9)$$

其中，术语 *OA* 是 遮挡感知 的缩写。

我们使用图2中的最后一个场景——遮挡共享——来演示这一概念。三条轨迹和两架无人机均已初始化，其中轨迹II起初不归任何无人机所有。如果沿着轨迹I飞行，无人机1将检测到唯一的遮挡时间 k_1。之后在时间 $k_2 > k_1$，轨迹III将进入同一遮挡区域。如果共享这一遮挡信息，我们预计轨迹III表现出遮挡感知行为的概率会更高。

遮挡共享的仿真分析：遮挡共享的感知行为以与图 6 中联合效应相同的方式进行了可视化。遮挡感知行为出现在两个时间区间：轨迹 I 上的 2.8 ∼ 3.6 秒，以及轨迹 III 上的 8.8 ∼ 9.6 秒。在第一次遮挡事件中，根据灰色区间内与轨迹 I 重叠的传感器视场平均数量，我们观察到其感知效果排序为： `先验、动态和无遮挡地图。由于轨迹 II 初始时未被覆盖，为了平衡无人机 1 和 2 带来的不确定性，覆盖范围有所增加：蓝线和红线在遮挡区域附近出现的无人机较少，以覆盖轨迹 II；绿线对轨迹 III 的感知次数较少。第一次时间区间内的行为实证结果与联合效应一致。然而，在第二个区间轨迹 III 被遮挡时，动态遮挡地图与 `先验遮挡地图的覆盖行为几乎相同。无遮挡地图的行为仍然在遮挡区域浪费感知资源，并且蓝线和红线曲线在三条轨迹上均相互重叠。动态遮挡地图与 `先验遮-挡地图的性能接近，表明在检测并共享遮挡信息后，遮挡感知有所增强。这验证了遮挡共享的定义。

V. 结论

本文定义了三种涌现协同行为：所有权变更、遮挡感知和遮挡共享。通过四种可能存在协同行为的 DTT 场景，¯所定义的涌现行为验证了

- 不同计算复杂度的规划算法的协作性：分散式部分可观测马尔可夫决策过程（Dec-POMDP）和 SMA-NBO；
- 与贪婪方法相比，非短视的求解方法所引发的协同行为有所增加；
- 动态遮挡地图在未知遮挡条件下的有效性；
- 基于动态遮挡地图探索共享范式的遮挡感知改进。

本文所展示的场景基于其简单的配置，展示了初步结果。然而，这已提供了充分的见解，可用于改进适用于更一般 SOO 几何结构的动态遮挡地图，并研究在线分布式遮挡融合方法。

参考文献

[1] Y. Lu、X. Xu、M. Ding、Z. Lu 和 T. Xiang，“一种面向全局遮挡的自监督单目视觉里程计方法”，载于 *AAAI* 人工智能会议论文集，第 35 卷第 3 期，2021 年，第 2260–2268 页。[2] J. Yao、J. Zhang、Z. Wang 和 L. Shao，“语义与上下文特征融合的鲁棒目标跟踪”，*IET* 图像处理，第 16 卷第 5 期，2022 年，第 1268–1279 页。[3] A. Ur-Rehman、S. M. Naqvi、L. Mihaylova 和 J. A. Chambers，“利用学习型变分贝叶斯聚类和社会力模型实现多目标跟踪与遮挡处理”，*IEEE* 信号处理汇刊，第 64 卷第 5 期，2015 年，第 1320–1335 页。[4] S. C.-H. Yang、D. M. Wolpert 和 M. Lengyel，“主动感知的理论视角”，行为科学当前观点，第 11 卷，2016 年，第 100–108 页。[5] T. Bandyopadhyay、N. Rong、M. Ang、D. Hsu 和 W. S. Lee，“不确定动态环境中的人员跟踪运动规划”，载于 人员检测与跟踪研讨会，*IEEE* 国际机器人与自动化会议，2009 年。

[6] R. Zou 和 S. Bhattacharya，“基于可见性的目标跟踪博弈中的最优追踪轨迹”，IEEE 机器人学汇刊，第 35 卷第 2 期，2018 年，第 449–465 页。[7] J. Gemerek、B. Fu、Y. Chen、Z. Liu、M. Zheng、D. van Wijk 和 S. Ferrari，“用于避免遮挡的定向传感器规划”，IEEE 机器人学汇刊，第 38 卷第 6 期，2022 年，第 3713–3733 页。[8] B. Jung 和 G. S. Sukhatme，“使用多个机器人跟踪目标：环境遮挡的影响”，自主机器人，第 13 卷第 3 期，2002 年，第 191–205 页。[9] N. Buckman、A. Pierson、S. Karaman 和 D. Rus，“生成可感知性驱动的协同主动运动规划轨迹”，载于 IEEE 国际机器人与自动化会议论文集，2020 年，第 3220–3226 页。[10] R. Tallamraju、E. Price、R. Ludwig、K. Karlapalem、H. H. B ̈ulthoff、M. J. Black 和 A. Ahmad，“基于主动感知的多飞行器编队控制”，IEEE 机器人与自动化快报，第 4 卷第 4 期，2019 年，第 4491–4498 页。[11] T. Li、L. W. Krakow 和 S. Gopalswamy，“利用多智能体滚动控制策略优化基于共识的多目标跟踪”，载于 IEEE 控制技术应用会议论文集，2021 年，第 131–137 页。[12] ——，“Sma-nbo：目标跟踪中采用名义信念状态优化的序贯多智能体规划”，载于 IEEE/RSJ 智能机器人与系统国际会议（IROS）论文集，2022 年，第 5861–5868 页。[13] M. Jakob、E. Semsch、D. Pavl ́ıcek 和 M. P ̌echoucek，“城市多个区域的具备遮挡感知能力的多无人机监视”，载于《智能体、交通与运输》研讨会，Citeseer，2010 年，第 59–66 页。[14] K. Hausman、G. Kahn、S. Patil、J. M ̈uller、K. Goldberg、P. Abbeel 和 G. S. Sukhatme，“具备遮挡感知能力的多机器人三维跟踪”，载于 IEEE/RSJ 智能机器人与系统国际会议（IROS）论文集，2016 年，第 1863–1870 页。[15] M. T. Spaan 、 T. S. Veiga 和 P. U. Lima，“ 利用 POMDP 在网络机器人系统中进行主动协同感知”，载于 IEEE/RSJ 智能机器人与系统国际会议（IROS）论文集，2010 年，第 4800–4805 页。[16] S. Ferrari 和 T. A. Wettergren，《信息驱动的规划与控制》。MIT 出版社，2021 年。[17] B. Mobedi 和 G. Nejat，“时间关键型城市搜索与救援任务中的三维主动感知”，IEEE/ASME 机电一体化汇刊，第 17 卷第 6 期，2011 年，第 1111–1119 页。[18] Y. Wu、Z. Zhao、S. Zhang、L. Yao、Y. Yang、T. Z. Fu 和 S. Winkler，“交互式多摄像机足球视频分析系统”，载于 ACM 多媒体国际会议论文集，2019 年，第 1047–1049 页。[19] D. A. Shell、L. Huang、A. T. Becker 和 J. M. O’Kane，“面向结构化叙事的协同事件观测规划”，载于 IEEE 国际机器人与自动化会议论文集，2019 年，第 7632–7638 页。[20] Y. Kantaros、B. Schlotfeldt、N. Atanasov 和 G. J. Pappas，“非短视的多机器人信息收集的渐近最优规划”。载于《机器人学：科学与系统》（RSS），2019 年，第 22–26 页。[21] G. L ́opez-Nicol ́as、M. Aranda 和 Y. Mezouar，“在运动与可见性约束下用于包围和跟踪目标的自适应多机器人编队规划”，IEEE 机器人学汇刊，第 36 卷第 1 期，2019 年，第 142–156 页。[22] S. A. Miller、Z. A. Harris 和 E. K. Chong，“通过名义信念状态优化实现自主无人机的协同引导”，载于美国控制会议论文集，2009 年，第 2811–2818 页。[23] V. Braitenberg，《车辆：合成心理学实验》。MIT 出版社，1986 年。[24] C. W. Reynolds，“群、畜群与鱼群：一种分布式行为模型”，载于 ACM SIGGRAPH 计算机图形学会议论文集，1987 年，第 25–34 页。[25] E. Pagello、A. D’Angelo、F. Montesello、F. Garelli 和 C. Ferrari，“通过隐式通信实现多机器人系统中的协同行为”，机器人与自主系统，第 29 卷第 1 期，1999 年，第 65–77 页。[26] P. Nalepka、R. W. Kallen、A. Chemero、E. Saltzman 和 M. J. Richardson，“驱赶那些羊：涌现式多智能体协调与行为模式切换”，《心理科学》，第 28 卷第 5 期，2017 年，第 630–650 页。[27] Y. Li、K. H. Johansson、J. M ̊artensson 和 D. P. Bertsekas，“用于确定性最优控制的数据驱动滚动规划”，载于 IEEE 决策与控制会议论文集，2021 年，第 2169–2176 页。[28] S. Ragi 和 E. K. Chong，“显式优化通信的无人机分散式引导控制”，《智能机器人系统期刊》，第 73 卷第 1 期，2014 年，第 811–822 页。