arXiv:2304.02075v2 [cs.RO] 9 Jan 2025

GUTS：面向多智能体主动搜索的广义不确定性感知汤普森采样

Nikhil Angad Bakshi[1], Tejus Gupta[1],Ramina Ghods[1] 和 Jeff Schneider[1]

摘要——快速开展灾害应急响应的机器人解决方案至关重要，尤其是在搜索区域对人类救援人员而言过于危险或范围过大时，这有助于最大限度地减少人员伤亡。我们将这一问题建模为一项异步多智能体主动搜索任务，其中每台机器人都旨在高效地在未知环境中搜寻感兴趣目标（OOI）。这种建模方式满足了搜索任务应侧重于快速找到感兴趣目标，而非全面覆盖搜索区域的要求。以往的方法未能准确建模感知不确定性，未能考虑植被或地形造成的遮挡，也未能兼顾对异构搜索队伍的需求以及应对硬件和通信故障的鲁棒性。我们提出广义不确定性感知汤普森采样（GUTS）算法，以解决这些问题；该算法适用于部署在异构多机器人系统上，用于在大型非结构化环境中进行主动搜索。仿真实验表明，GUTS 的表现始终优于并行化汤普森采样和穷举搜索等现有方法，在80% 的运行中找到了全部感兴趣目标。相比之下，现有方法在不到40% 的运行中找到了全部感兴趣目标。我们在非结构化环境中使用多机器人系统开展了实地测试，搜索区域面积为 $\approx 75{,}000\ m^2$。我们的系统展现出应对多种故障模式的鲁棒性，在每次实地测试中都找到了全部感兴趣目标（在可行的情况下），并且表现显著优于基线方法。

I. 引言

多机器人系统已越来越多地应用于大规模搜救行动 [2], [3], [4], [5]。此类应急行动时间紧迫，因此搜索效率至关重要。人类操作员可以远程控制少量车辆，但无法高效协调规模更大的团队。这促使人们需要开发去中心化的多机器人系统，使其能够在大范围、无结构的陆地环境中高效开展主动搜索。

主动搜索是指在适应新感知信息的同时，通过高效的序贯数据采集决策来识别位置分散的感兴趣目标（OOIs） [1], [7]。主动搜索的基本框架包括维护感兴趣目标在各个位置上的后验概率分布，并优化信息搜寻行动。这些行动必须在勘察未探索区域（探索）与确认疑似感兴趣目标（利用）之间取得平衡 [21], [22], [23]。

尽管这一基本框架简单明了，但要在真实搜救场景中将这些算法部署于多机器人系统，还需解决若干额外挑战。感兴趣目标可能被遮挡，无法从俯视或侧视视角观察；因此，需要结合地面和空中。

[1]N. A. Bakshi、T. Gupta、R. Ghods 和 J. Schneider 均供职于卡内基梅隆大学计算机科学学院机器人研究所，宾夕法尼亚州匹兹堡 15213 {nabakshi, tejusg, rghods, schneide}@cs.cmu.edu

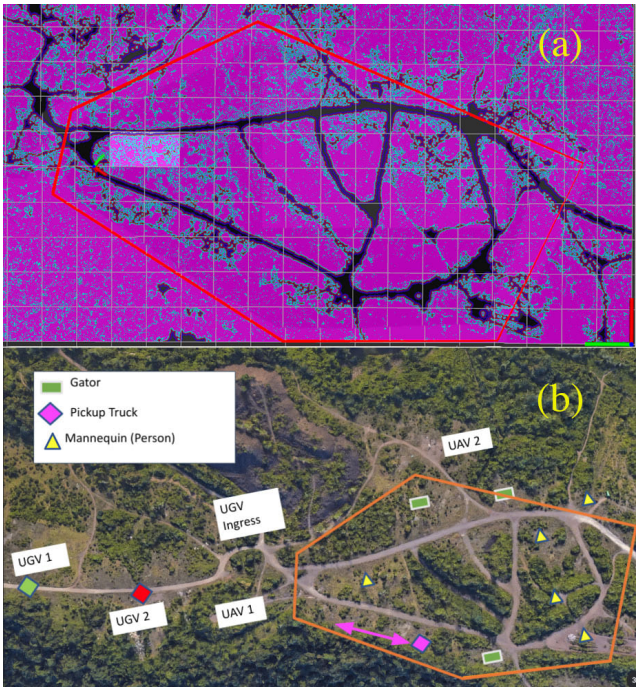


图 1：宾夕法尼亚州匹兹堡的测试场地。(a) 显示遍历代价地图和搜索多边形（紫色表示高代价，黑色表示低代价）；(b) 显示一张俯视图，其中标出了最常见的感兴趣物体位置和机器人的出发区域。多边形内的面积约为 75,000 平方米。

搜索区域车辆对于彻底勘查此类搜索区域至关重要。这要求搜索算法在规划搜索动作时，推理遮挡、可视域约束，以及每个机器人的感知模型和导航约束。类似地，尽管多机器人团队理论上可以划分搜索区域以提高探索效率 [25]，但在去中心化系统中生成此类划分颇具挑战。尽管我们希望机器人尽可能多地共享感知信息，以实现有效协同，但不能依赖稳定的通信。因此，我们要求系统具备异步特性，并能够利用部分信息做出明智决策。最后，搜索区域中的感兴趣目标数量通常未知，且感知动作存在噪声 [30],[31], [32]。以往的搜索方法采用了过度简化的假设 [19], [6]，且未能充分应对这些挑战，因此难以将其部署到真实的搜索场景中。

我们提出一种新颖的搜索算法——广义不确定性感知汤普森采样（GUTS），该算法遵循并行化汤普森采样框架 [1], [7], [8]。我们提出了一种改进的奖励函数，以现实地建模

无人机和无人地面车辆中的感知噪声，并对其进行计算优化，使其能够在大型环境中运行。

我们设计 GUTS，使其适用于任何非结构化搜索区域，不对感兴趣目标的数量作出假设，并能抵御通信中断和硬件故障。例如，即使部分机器人发生故障并停止响应，我们的搜索团队仍能继续运行。

我们在宾夕法尼亚州匹兹堡的大型非结构化自然环境中进行了实地测试，并开展了仿真测试，展示了由无人地面车辆和无人机组成的搜索系统，如图 1 所示。结果表明，我们的系统具有鲁棒性，且搜索性能优于现有方法。据我们所知，这是首个以这一规模构建的异构多机器人系统，用于异步多智能体主动搜索；该系统能够抵御通信中断和机器人故障，且无需人工指挥或手动划分搜索区域即可运行。

II. 相关工作

以往的多机器人搜索系统研究针对的场景更为专门（例如，同质团队、单智能体搜索），且采用了更强的假设（通信完美、同方差噪声模型）。例如，基于覆盖的搜索方法 [10], [11], [12], [24] 会穷尽式地遍历搜索空间。这些方法假设搜索区域已有地图，且无法在线适应新信息。对 [27] 等基于前沿的方法也存在类似的批评。

贝叶斯主动搜索方法 [30], [31] [32] 能够在单个智能体搜索稀疏信号时准确地对观测不确定性进行建模。然而，这些方法的搜索策略具有确定性，因此在不进行集中式控制的情况下，很难将其扩展到多智能体场景。近期研究将主动搜索问题建模为部分可观测马尔可夫决策过程，并使用强化学习在仿真环境中训练搜索策略 [22], [21], [23], [38], [28]。由于样本复杂度高且仿真环境不切实际，这些方法通常难以部署到机器人系统中。

Ghassemi 等人 [6] 采用信息论方法开展异步多智能体主动搜索，但其研究对象是一个同质群体，并假设通信完美且采用同方差噪声模型。Salman 等人 [19] 生成轨迹，使得在某一区域花费的时间与该区域存在感兴趣目标的似然相对应。然而，他们的方法不纳入新的观测，并依赖于完美通信以及关于感兴趣目标位置的强统一先验。

Dec-MCTS [26] 是一种多智能体主动感知算法，已展现出有前景的仿真结果。Dec-MCTS 是蒙特卡洛树搜索的去中心化版本，所有机器人异步优化各自的轨迹，并通过共享部分搜索树实现协调，但不对实际感知流程中的不确定性进行建模。

噪声感知汤普森采样（NATS） [1] 是一种异步多智能体主动搜索算法，能够处理异方差噪声模型，并已展现出良好的仿真结果。NATS 使用 Myopic Posterior Sampling（MPS） [9] 确保智能体能够在不依赖集中式协调的情况下采取多样化的感知动作。汤普森采样（即 MPS） [15]是一种在线优化方法，它通过假设从后验分布中抽取的样本就是真实的世界状态 [16]，最大化预期奖励，从而平衡探索与利用。这种随机性是使其非常适用于通信不可靠的异步多智能体场景的关键特性，因为它能促进动作选择的多样性。





图 2：左：RecBot 无人地面车辆。右：Lil Hexy 无人机。

III. 贡献

- 我们提出广义不确定性感知汤普森采样（GUTS）算法。该算法采用改进的奖励设计，更准确地对感知噪声进行建模，并在计算效率方面优于先前的并行化汤普森采样方法 [1], [7]。
- 我们展示了一套可运行的异构机器人系统，能够在大型非结构化自然环境中开展无需中央规划器的异步多智能体主动搜索。我们设计的多智能体搜索算法能够在无需人工干预的情况下，抵御搜索过程中机器人损失、观测噪声以及通信中断。
- 通过实地和仿真实验，我们证明，与 NATS 和基于覆盖的搜索等现有搜索方法相比，本方法具有更高的搜索成功率和搜索效率。

IV. 系统描述

本节介绍我们用于评估搜索算法的多机器人系统。对于无人地面车辆（UGV），我们使用 RecBot——一辆经过改装、配备线控自动驾驶套件的 John Deere E-Gator 多用途车辆（图 2 左）。对于无人机（UAV），我们使用 Lil Hexy [1] 机架与 PixHawk 飞行控制器 [2] 组成的基础平台（图 2 右）。我们为该平台加装了定制相机载荷和额外的机载计算设备，即 NVIDIA Jetson AGX Xavier。

下面总结理解本文工作所需的自主系统技术栈组件。

- **相机视场：** 无人地面车辆上的传感器舱具有37° 垂直视场和 192° 水平视场（图 3-左）。传感器模块由五个 RGB相机构成，每个相机的分辨率为 12

[1]https://vscl.tamu.edu/vehicles/little-hexy/ [2] https://pixhawk.org/

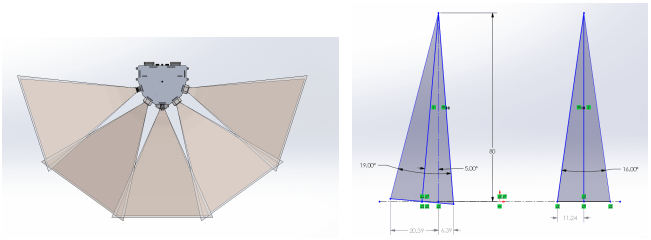


图 3：左：无人地面车辆相机视场的俯视图。水平跨度为 192°，垂直跨度为 37°。右：无人机相机视场。跨度为 19° × 16°。在 80 米的飞行高度下，这使得地面上形成一个约为 27.8米 ×22.5米 的可见矩形区域。

MP，以便在数百米外准确检测感兴趣目标。无人机上安装了一个配备单个 5 MP 分辨率 RGB相机的传感器舱，相机朝下，并相对于垂直方向倾斜 5° 角，倾斜方向与无人机的飞行方向一致（图 3-右）。从 80m 的飞行高度拍摄时，该相机的 19° × 16° 视场可覆盖 $27.8m \times 22.5m$ 的区域。

- **目标检测器：** 对于无人地面车辆，我们使用广泛采用的 YOLOv3 神经网络架构的定制版本 [13]，在输入图像分辨率极高的情况下，利用 NVIDIA RTX 6000 图形处理器实现快速推理。对于无人机检测器，由于计算能力有限，因此我们使用计算开销较低的模型 [14]，并在 NVIDIA Jetson AGX Xavier 上运行。

- **目标跟踪：** 感知模块使用卡尔曼滤波器 [34] 跟踪可能的感兴趣目标。我们在第六节中详细描述感知模型。例如，与方位方向相比，地面车辆检测到的感兴趣物体位置估计在沿物体距离方向上的不确定性更高。这由卡尔曼滤波器返回的跟踪轨迹所关联的协方差表示。此外，这种不确定性会随着物体与观测者之间的距离增加而增大。

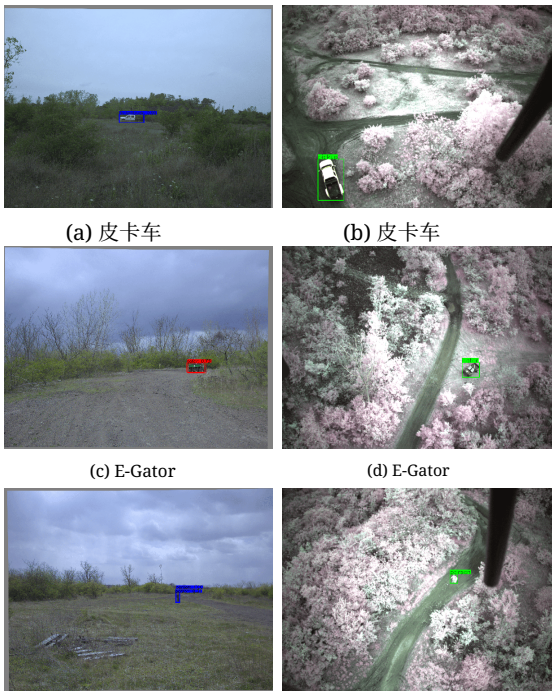
- **无人地面车辆导航：** 我们采用SBPL规划器[3]，并结合局部前视控制器，为本文介绍的搜索模块提供的航路点规划安全路径。我们利用机器人搭载传感器感知更新离线全局地图，以生成遍历代价地图。

- **感兴趣目标：** 感兴趣目标包括人员（人体模型）、一辆 John Deere E-Gator 或一辆皮卡车。图 4 展示了无人地面车辆和无人机相机系统拍摄的示例画面，其中成功检测到了目标。

V. 问题描述

我们将搜索区域建模为网格，网格单元尺寸为 30m x 30m。我们可以获取用于评估地面机器人可通行性的代价地图。图 1 展示了宾夕法尼亚州匹兹堡实地测试地点的俯视图和代价地图。该代价地图是实地测试前数月通过无人机航拍生成的。此后，地图的部分区域发生了变化，但我们搭载在机器人上的感知和建图系统能够识别道路受阻等变化。鉴于卫星影像十分普及，假设能够获取搜索区域的大致地图信息是合理的。

https://github.com/sbpl/sbpl



(a) 皮卡车 (b) 皮卡车 (c) E-Gator (d) E-Gator (e) 人 (f) 人

图 4：目标检测器：定性结果。我们展示了无人地面车辆（左）和无人机（右）的图像及感兴趣物体检测结果

如图 1 所示，地面车辆无法到达或观察若干区域。此外，地图的大片区域覆盖着茂密植被，可能会遮挡空中飞行器视野中隐藏在植被下的任何感兴趣目标。因此，需要通过异构协同才能完整探索搜索区域。

我们将主动搜索问题建模如下：

- 我们为所有机器人提供相同的搜索区域。（示例见图1中的红色多边形）。
- 感兴趣目标以稀疏方式分布且保持静止。需要以尽可能快的速度并以高确定性定位它们。
- 每个机器人必须在机载设备上规划其下一步数据采集行动，即不存在中央规划器。
- 机器人可以相互通信其位置和观测结果；但是，该算法应能在从完全没有通信到完美通信的整个通信范围内运行。

我们注意到，该目标并非覆盖整个搜索空间，而是具体地以尽可能快的速度定位所有感兴趣目标。

更正式地说，我们使用稀疏矩阵$\mathbf{B} \in \mathbb{R}^{M_1 \times M^2}$，在二维网格表示中表示感兴趣目标的位置。令 $\beta \in \mathbb{R}^M$ 为矩阵$\mathbf{B}$的展平版本，其中$M = M_1 M_2$。该向量中只有少数位置有值，对应于感兴趣目标的位置为1，其余位置为0。我们通过以下方式对观测中的不确定性进行建模：

$$\mathbf{y}_t = \text{clip}(\mathbf{X}_t \beta \pm \mathbf{n}_t, 0, 1) \text{ with } \mathbf{n}_t \sim \mathcal{N}^+(0, \Sigma_t) \quad (1)$$

其中，$\mathbf{X}_t \in \mathbb{R}^{Q \times M}$ 表示感知矩阵，其每一行在$\mathbf{X}_t$ 中都是一个独热向量，用于指示机器人在时间步$\mathbf{t}$可见的一个网格

单元，而Q表示机器人可见的网格单元总数。$\mathbf{y}_t \in \mathbb{R}^{Q\times 1}$是包含加性深度感知噪声向量$\mathbf{n}_t \in \mathbb{R}^{Q\times 1}$的结果观测。这种深度感知噪声体现了观测不确定性会随距离机器人的距离增加而增大的直觉；我们将噪声协方差矩阵的对角元素 Σ_t 建模为随观测网格单元与机器人之间的距离单调增加。Σ_t 的非对角元素为零。噪声从正半高斯分布 $\mathcal{N}^+(0,\Sigma_t)$中采样；对于不含感兴趣目标的网格单元，将其加上；对于含有感兴趣目标的网格单元，则将其减去。观测y_t 被截断在 0 到 1 之间。

令$\mathbf{D}_{jt}$表示时刻t机器人j可用的观测数据。$\mathbf{D}_{jt}$由机器人j收集的（$\mathbf{X}_t$,$\mathbf{y}_t$）对，以及其他机器人传送给机器人j 的数据组成。设所有智能体执行的感知动作总数为T。我们的主要目标是在每个时刻t ，根据$\mathbf{D}_t^j$ 依次优化下一个感知动作$\mathbf{X}_{t+1}$ ，以尽可能少的测量T 恢复稀疏信号 β 。每个机器人都基于自身的部分数据集$\mathbf{D}_t^j$，以去中心化方式优化这一目标。

VI. 广义不确定性感知汤普森采样（GUTS）

本节介绍 GUTS 算法。决策框架遵循 [1]。每个机器人j 都会基于其部分数据集$\mathbf{D}_t^j$，异步估计感兴趣物体位置的后验分布。在动作选择阶段，每个机器人从该后验分布中生成一个样本，并针对这组采样得到的感兴趣物体位置优化奖励函数。第 VI-A 节和第 VI-B 节介绍了文献 [1]中提出的噪声感知汤普森采样（NATS）算法的这两个步骤。我们在 GUTS 中采用第 VI-A 节所述的后验分布计算步骤。

在第 VI-C 节及后续内容中，我们将介绍用于提升搜索性能的新奖励函数、适用于无人地面车辆和无人机的深度感知噪声模型，以及用于优化代码、使其能在大型环境中运行的加速技术。

A. 计算后验分布

每个机器人都假设向量 β 的每个分量服从零均值高斯先验，即 $\mathrm{p}_0(\beta_m) = \mathcal{N}(0,\gamma_m)$。方差$\Gamma$ =diag([$\gamma_1 \ldots \gamma_M$]) 是通过数据估计的隐变量。我们遵循 [17], [18] ，并对 γ_m 使用共轭逆伽马先验以保证稀疏性，即 $\mathrm{p}(\gamma_m)$ =IG(a_m, b_m) =bammr(am) $\gamma^{(-a_m-1)}$)m $\mathrm{e}^{-(b_m/\gamma}$m) $\forall$m $\in \{1 \ldots \mathrm{M}\}$。我们使用期望最大化 [40]，根据机器人 j 的数据 D jt 来估计 β 的后验分布。

我们可以分别写出 E 步（估计 $p(\beta|\mathbf{D}_{jt}\Gamma) = \mathcal{N}(\mu,$ **V**)) 和 M 步（计算 $\max_\Gamma p(\mathbf{D}_{jt}|\beta,\ \Gamma)$）的解析表达式：

$$\mathbf{V} = (\Gamma^{-1} + \mathbf{X}^T\Sigma\mathbf{X})^{-1}; \mu = \mathbf{V}\mathbf{X}^T\Sigma\mathbf{y} \tag{2}$$

$$\gamma_m = ([\mathbf{V}]_{mm} + [\mu]_m^2 + 2b_m)/(1+2a_m) \tag{3}$$

其中，$\mathbf{X}$和$\mathbf{y}$ 由$\mathbf{D}_{jt}$中的所有测量值（$\mathbf{X}_t$，$\mathbf{y}_t$）纵向堆叠而成，Σ是由其对应的深度感知噪声方差构成的对角矩阵。

每个机器人都使用其部分数据集D_{jt}在本地估计$p(\beta|\mathbf{D}_{jt})$ $= \mathcal{N}(\mu,\ \mathbf{V})$。为清晰起见，我们在公式（2）和（3）中省略索引$j$。我们将$a_m = 0.1$ 和$b_m = 1$的值设为在 [1]中发现有效的取值。最后，智能体j从后验分布 $\tilde{\beta} \sim p(\beta|\mathbf{D}_t^j)$中采样。

B. 选择下一个感知动作

每个机器人都假设采样得到的感兴趣物体位置集合 β 是正确的，并据此选择下一个感知动作 $\mathbf{X}_{t+1}$ 。具体而言，令 $\hat{}\ \beta(\mathbf{D}_{jt}\cup(\mathbf{X}_{t+1},\ \mathbf{y}_{t+1}))$ 表示利用所有可用测量数据 $\mathbf{D}_{jt}$ 以及下一个候选测量 $(\mathbf{X}_{t+1},\ \mathbf{y}_{t+1})$ 对参数 β 的期望估计值。于是，用于评估每个可能感知动作的奖励函数定义如下：

$$\mathscr{R}(\tilde{\beta},\mathbf{D}_t^j,\mathbf{X}_{t+1}) = -\mathbb{E}_{\mathbf{y}_{t+1}|\mathbf{X}_{t+1},\tilde{\beta}}[||\tilde{\beta} - \hat{\beta}(\mathbf{D}_t^j\cup(\mathbf{X}_{t+1},\mathbf{y}_{t+1}))||_2^2] \tag{4}$$

我们选择使该奖励最大化的感知动作 $\mathbf{X}_{t+1}$ 。由于对 β 的采样会使奖励函数具有随机性~，因此机器人选择的搜索动作具有多样性。

C. GUTS 奖励函数

在动作选择过程中，每个机器人通过最大化公式（4）中的奖励函数来优化感知位置。我们发现，在运行初期，该奖励函数会导致搜索性能较差。在回合开始时，关于感兴趣物体位置的后验信念非常不确定，所有可能的下一个航路点几乎都不可能找到物体。因此，该奖励函数会导致过度探索的行为。

我们将奖励函数修改为

$$\mathscr{R}(\tilde{\beta},\mathbf{D}_t^j,\mathbf{X}_{t+1}) = -\mathbb{E}_{\mathbf{y}_{t+1}|\mathbf{X}_{t+1},\tilde{\beta}}[||\tilde{\beta} - \hat{\beta}(\mathbf{D}_t^j\cup(\mathbf{X}_{t+1},\mathbf{y}_{t+1}))||_2^2] - \lambda \times I(\tilde{\beta},\hat{\beta}) \tag{5}$$

其中，$\lambda(=0.01)$ 是一个超参数：如果估计的 β 与当前步骤 β 的样本 ~ 之间没有较高似然的共同条目，它就会降低搜索位置的奖励。此修改增强了奖励的区分能力，并鼓励在搜索过程早期采取利用性动作。形式上，设 k 和 k 分别为 β 和 β 中非零条目的数量，则指示函数 $I(.)$ 定义如下：

$$I(\tilde{\beta},\hat{\beta}) = \begin{cases} 0, \text{ if any matches between top } \frac{\hat{k}}{2} \text{ entries in } \hat{\beta}, \\ \quad \text{and top } \frac{\tilde{k}}{2} \text{ entries in } \tilde{\beta} \\ 1, \text{ otherwise} \end{cases}$$

实验表明，这种修改后的奖励函数能够提高搜索效率。

D. 无人机感知动作模型

每架无人机的飞行高度固定在 70米 − 90米，其下方地面的视场约为 30米 × 30米 。我们将飞行路径视为直线，并在计算奖励时对飞往最终航路点途中经过的网格单元进行积分。

E. UGV 感知动作模型

我们将 UGV 的视场建模为其正前方的 30米 $\times$ 60米空间。在 GUTS 网格表示中，这将是沿 UGV 方位方向的两个网格单元。与图 3 所示的完整视场相比，我们对视场进行保守近似。高估视场将是灾难性的，因为 GUTS 随后会错误地认为搜索区域的某些部分已经被搜索过，从而永远无法在该区域发现感兴趣目标。请注意，即使超出建模的视场，感知模块仍会报告阳性检测，因此不会丢失信息。

我们在计算无人地面车辆的奖励函数时，不考虑候选感知动作 $\mathbf{X}_t$ 途中可能出现的感兴趣目标，因为针对每种可能的感知动作 $\mathbf{X}_t$ 计算并对轨迹进行积分的代价很高。尽管沿途观测未被明确纳入计算，每个机器人仍会记录并共享前往航路点途中的观测。未来工作将研究如何整合无人地面车辆的轨迹级信息增益。

F. 改进的深度感知噪声建模

本节介绍我们用于感兴趣物体观测的异方差噪声模型。NATS [1] 使用对角协方差矩阵对这种噪声建模，其中对角元素与被观测网格单元和机器人之间的距离成正比。直观而言，这表示我们对在所考虑网格单元处检测到的感兴趣物体是否存在的置信度。然而，对于基于学习的感知系统而言，这一量实际上无法获取。

我们将物体检测器的检测置信度与通过卡尔曼滤波器获得的位置不确定性相结合，以估计这一置信度。我们采用以下启发式方法来近似正向观测中的存在不确定性：

$$\sigma^2 = min\left(0.5, \frac{\text{Vol. of OOI loc uncertainty ellipsoid } (m^3)}{\text{OOI detection confidence} \times 1000}\right)$$

这个公式基于这样的直觉：卡尔曼滤波器得到的位置不确定性也隐含编码了存在不确定性。对于负向观测，我们重用 [1] 中采用的原始深度感知噪声模型。

G. 加速 *GUTS*

我们通过一个简单的示例来说明加速的幅度。对于面积为 0.7 平方千米的区域，原始 NATS 实现 [1] 计算第一个航路点耗时超过 60 分钟。经过以下三项改进后，我们将航路点选择时间缩短至不到两分钟：

- GUTS 需要为候选感知动作 $\mathbf{V}$ 在式（2）中反复计算$\mathbf{X}_{t+1}$。矩阵求逆步骤开销很大。我们利用 $(\Gamma^{-1} + \mathbf{X}_T\ \Sigma\mathbf{X})^{-1}$ 的对角结构，只需对对角元素求逆即可加速这一过程。
- 此外，(2) 中的 $\mathbf{X}_T\ \Sigma\mathbf{X}$ 项可以按如下方式进行加性计算：

$$\begin{bmatrix}\mathbf{X}_1\\ \mathbf{X}_2\end{bmatrix}^T \begin{bmatrix}\Sigma_1 & 0\\ 0 & \Sigma_2\end{bmatrix} \begin{bmatrix}\mathbf{X}_1\\ \mathbf{X}_2\end{bmatrix} = \mathbf{X}_1^T\Sigma_1\mathbf{X}_1 + \mathbf{X}_2^T\Sigma_2\mathbf{X}_2$$

这利用了不同网格单元中感知噪声的独立性，并在运行后期$\mathbf{X}$ 维度较高时显著加快计算速度。

- 对于式（2）中的 μ，我们利用感知矩阵X 的稀疏性来加速计算 [37]。

由于无人机的计算资源有限，我们增加了一个采样参数，用于对无人机可能采取的感知动作进行下采样，以便按需调整规划时间。通常，我们测试的下采样比例范围为 1% 到 10%。

VII. 实验与结果

我们的实验旨在展示，与现有搜索方法 NATS 和基于覆盖的搜索相比，我们提出的算法 GUTS 能够提高搜索效率。由于原始实现 [1] 对于我们实验的规模而言运行过慢，因此我们在实验中采用了经过计算优化的 NATS 版本（类似于第 VI-G 节）。基线方法采用贪心策略，在搜索区域中尚未访问的部分选择下一个航路点。我们在逼真的多机器人 Gazebo [39] 仿真环境中评估这些方法。我们的仿真使用了实地测试地点的预先绘制的代价地图，并准确模拟了机器人的感知、轨迹规划和遍历。

我们还旨在验证 GUTS 能否在异构多机器人团队中实时运行，并能否在通信中断、硬件故障和含噪检测结果的情况下保持稳健。我们在自然的非结构化环境中开展实地测试，使用地面和空中车辆组成的团队，搜索区域面积为 $\sim 75,000$米^2。

A. 测试设置

图 1b 展示了无人地面车辆和无人机的出发位置、每台机器人的搜索区域以及最常见的感兴趣物体位置。机器人进入搜索多边形后，我们便开始评估其搜索性能。

B. 评估指标

我们的主要评估指标是搜索成功率，定义为在规定的时间预算内找到所有感兴趣目标的运行次数所占的比例。

对于仿真实验，我们通过绘制每种算法的召回率与运行时间关系图来衡量搜索效率。对于实地测试，我们计算每找到一个感兴趣物体所需的感知动作次数。我们希望将每种搜索算法的性能与实地测试中的地形状况等物理因素区分开来，并以决策次数而非总耗时来衡量召回率。尽管如此，我们也报告每次运行的实际运行时间，而且每次运行的总搜索时间始终不到 30 分钟。

C. 仿真结果

我们在图 5 中比较了 GUTS 与 NATS 和基于覆盖的搜索的搜索效率。图中展示了一台无人地面车辆（中图）和两台无人地面车辆（右图）在五个感兴趣目标、1000 秒（$\approx$ 17分钟）运行时间预算下的结果，分别取 20 次和 10 次运行的平均值。

表 I：实地测试结果：我们报告每次运行的搜索参数（团队规模、总搜索面积、感兴趣目标总数、所用算法）及评估指标（运行时间、发现的感兴趣目标数量和搜索效率 T/C）。我们观察到，GUTS 的搜索效率（T/C）优于基于覆盖的规划器。我们还记录了发生通信中断（覆盖基线）、硬件故障（硬件故障）以及探测器问题（噪声探测器 = 噪声探测器）的运行情况。

Run	团队规模（J）	搜索区域（m^2）	运行时间（s）	发现的感兴趣目标数量（C）	感兴趣目标总数（M）	T	T/C	算法	备注
1	1 架无人机	~75，000	1600	4	7	27	6.75	覆盖范围	-
2	1 架无人机	~75，000	1500	5	8	28	5.6	覆盖范围	-
3	1 架无人机	~75,000	1000	2	8	8	4	5% GUTS	-
4	1 架无人机	~75,000	1600	4	8	16	4	5% GUTS	-
5	1 架无人机	~75,000	500	3	4	5	1.67	GUTS	-
6	2 辆无人地面车辆	~75，000	1000（平均值）	3	3	3	1	GUTS	ND
7	2 辆无人地面车辆	~75，000	600（平均）	6	6	6	1	GUTS	覆盖基线，噪声探测器
8	2 辆无人地面车辆，1 架无人机	~75，000	1200（平均）	6	6	13	2.167	GUTS	覆盖基线、硬件故障、噪声探测

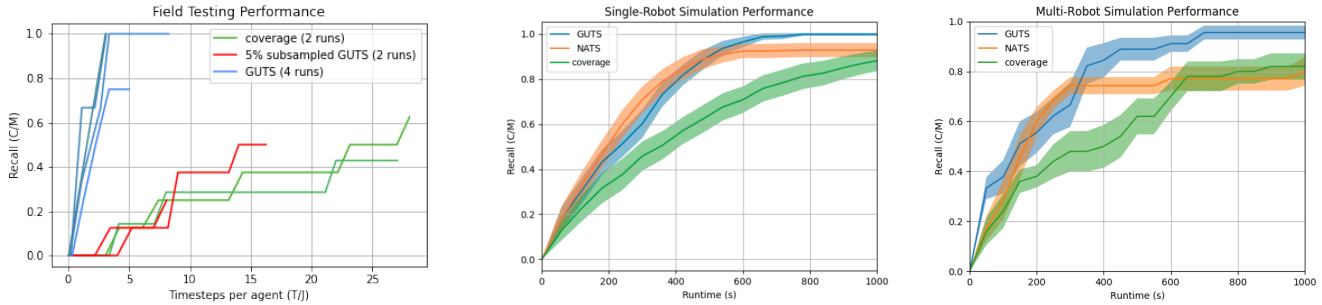


图 5：搜索效率比较。（左）我们展示表 I 中每次运行的搜索效率。图中绘制了召回率与每台机器人的感知动作次数之间的关系。该图表明，在我们的系统中，GUTS 的表现优于覆盖基线方法。我们还观察到，对全部航路点进行子采样，在计算效率和性能之间实现了良好平衡，并且在无人机上的表现略优于覆盖基线方法。（中和右）我们比较了单机器人和多机器人仿真中不同搜索算法的召回率与运行时间。我们再次观察到，GUTS 的表现优于 NATS 和基于覆盖的搜索。

可以看出，GUTS 明显优于我们的基线方法：NATS 和基于覆盖的搜索；它也是唯一一种能够在时间预算内稳定找回所有感兴趣目标的搜索算法。在多机器人场景下，GUTS 的成功率为 80%（即在 80% 的运行中找回所有感兴趣目标），而 NATS 和基于覆盖的搜索的成功率分别为 40% 和 30%。

由于 VI-C 中所述对奖励函数的修改，GUTS 的表现优于 NATS：该修改优先确认疑似感兴趣物体，而非采取探索行为。因此，GUTS 起步较慢，但在运行后期超过了 NATS。基于覆盖的搜索表现不佳，因为它会穷尽式地遍历搜索区域，而不是优化感知动作以识别感兴趣物体。

D. 实地测试结果

表 I 汇报了实地测试的详细运行参数和评估指标。我们报告了 GUTS 的 6 次运行结果，以及基于覆盖范围的基线方法的 2 次运行结果。图 5 左图绘制了每次运行中每个智能体所需的感知动作次数（T/J）与召回率之间的关系。我们观察到，与基于覆盖范围的基线方法相比，GUTS算法的搜索行为明显更加高效。

我们还观察到，在所有可能的感知动作中仅对 5% 的随机子样本运行的 5% GUTS，其表现也略优于覆盖范围。我们注意到，在出现通信中断或硬件故障的运行中，性能并未下降。在仅使用无人机的运行中，搜索团队有时无法找回所有感兴趣目标，因为有些目标只能由地面车辆找回。这表明需要同时使用无人机和无人地面车辆来彻底搜索该搜索区域。

VIII. 结论

我们提出了一种用于高效多智能体主动搜索的新算法——广义不确定性感知汤普森采样（GUTS）。GUTS 使用改进的奖励函数，能够真实地模拟感知噪声，并经过计算优化，可在大型环境中运行。我们通过仿真实验表明，与最先进的方法相比，GUTS 能够更快速、更稳定地找回所有物体。我们还通过在存在硬件故障和通信故障的大型非结构化环境中开展实地测试，验证了 GUTS 的搜索效率和鲁棒性。

未来的工作将重点关注动态目标建模，并结合路径级信息增益和地形信息来选择感知动作，从而进一步提升无人地面车辆的搜索性能。

我们认为，在危险的大规模搜救任务中更多地采用多机器人团队，有望拯救生命并减少人力投入。我们希望本研究能够推动人们在更贴近现实的搜索环境中开展研究，并加快其在实际搜救任务中的应用。

IX. 致谢

本材料基于美国陆军研究办公室和美国陆军未来司令部资助的工作，合同编号为 W911NF-20-D-0002。作者谨向 Conor Igoe、Arundhati Banerjee、Herman Herman、Jesse Holdaway、Prasanna Kannappan、Luis E. Navarro-Serment、Matthew Schnur、Wennie Tabib 和 Vamsavardan Vemuru 致谢，感谢他们所作的贡献。

参考文献

[1] R. Ghods、W. J. Durkin 和 J. Schneider，《采用真实深度感知噪声模型的多智能体主动搜索》，IEEE国际机器人与自动化会议（ICRA），2021

[2] C. Mouradian、J. Sahoo、R. H. Glitho、M. J. Morrow 和 P. A. Polakos，《面向大规模自然灾害中多机器人任务分配的联盟形成算法》，2017年第13届国际无线通信与移动计算会议（IWCMC），2017年，第1909–1914页，doi: 10.1109/IWCMC.2017.7986575。[3] D. S. Drew，《用于搜救应用的多智能体系统》，Current Robotics Reports，第2卷，第2期，第189–200页，2021年。[4] K. Nagatani 等，《用于搜救任务的多机器人探索：RoboCupRescue 2009中的地图构建报告》，Journal of Field Robotics，第28卷，第3期，第373–387页，2011年。[5] Baxter, J. 等，《多机器人搜救：一种基于人工势场的方法》，自主机器人与智能体，第9–16页，2007年。[6] P. Ghassemi 和 S. Chowdhury，《面向群体机器人搜索的平衡探索与利用分布式信息路径规划》，第1卷：第39届计算机与信息工程会议，2019年8月。[7] R. Ghods、A. Banerjee 和 J. Schneider，《针对稀疏信号的分布式多智能体主动搜索》，人工智能中的不确定性，第696–706页，2021年。[8] K. Kandasamy、A. Krishnamurthy、J. Schneider 和 B. Poczos，《通过汤普森采样实现并行贝叶斯优化》，载于第二十一届国际人工智能与统计学会议论文集，2018年4月9日至11日，第84卷，第133–142页。[9] K. Kandasamy、W. Neiswanger、R. Zhang、A. Krishnamurthy、J. Schneider 和 B. Poczos，《面向自适应目标导向实验设计的短视后验采样》，载于第36届国际机器学习会议论文集，2019年6月9日至15日，第97卷，第3222–3232页。[10] L. Lin 和 M. A. Goodrich，《用于荒野搜救的无人机智能路径规划》，载于2009年IEEE/RSJ智能机器人与系统国际会议。电气电子工程师学会，2009年，第709–714页。[11] S.-Y. Chien、H. Wang 和 M. Lewis，《机器人团队搜救中的人工与算法路径规划》，载于人因与工效学学会年会，第54卷，第4期。SAGE，2010年。[12] A. Ryan 和 J. K. Hedrick，《用于无人机辅助搜救的模式切换路径规划器》，载于决策与控制会议。电气电子工程师学会，2005年，第1471–1476页。[13] J. Redmon 和 A. Farhadi，《YOLOv3：渐进式改进》，arXiv预印本 arXiv:1804.02767，2018年。[14] W. Ke、T. Zhang、Z. Huang、Q. Ye、J. Liu 和 D. Huang，《用于视觉目标检测的多锚点学习》，载于2020年IEEE/CVF计算机视觉与模式识别会议（CVPR），2020年，第10 203–10 212页。[15] W. R. Thompson，《根据两组样本的证据，一个未知概率大于另一个未知概率的可能性》，Biometrika，第25卷，第3/4期，第285–294页，1933年。[16] D. J. Russo、B. Van Roy、A. Kazerouni、I. Osband、Z. Wen 等，《汤普森采样教程》，机器学习基础与趋势®，第11卷，第1期，第1–96页，2018年。[17] M. E. Tipping，《稀疏贝叶斯学习与相关向量机》，机器学习研究期刊，第1卷，6月号，2001年。[18] D. P. Wipf 和 B. D. Rao，《用于基选择的稀疏贝叶斯学习》，IEEE信号处理汇刊，第52卷，第8期，2004年。[19] H. Salman、E. Ayvali 和 H. Choset，《具有避障功能的多智能体遍历覆盖》，国际自动规划与调度会议论文集，第27卷，第242–249页，2017年。[20] E. Ayvali、H. Salman 和 H. Choset，《受限环境下基于随机轨迹优化的遍历覆盖》，arXiv [ ], cs.RO。2017

[21] X. Ye、Z. Lin、H. Li、S. Zheng 和 Y. Yang，《主动目标感知器：面向移动机器人目标搜索的识别引导策略学习》，载于2018年IEEE/RSJ智能机器人与系统国际会议（IROS），2018年，第6857–6863页。[22] Y. Zhu 等，《利用深度强化学习在室内场景中进行目标驱动视觉导航》，载于2017年IEEE国际机器人与自动化会议（ICRA），2017年，第3357–3364页。[23] N. Xu，《移动机器人利用强化学习进行主动目标搜索》，载于2021年第二届国际计算与数据科学会议（CDS），2021年，第296–300页。

[24] A. Sadeghi 和 S. L. Smith，“通过去中心化大邻域搜索实现异构任务分配与排序”，无人系统，第 05 卷，第 02 期，第 79–95 页，2017年。[25] D. W. Schuldt 和 J. A. Kurucar，“在无障碍环境中高效划分空间以执行多无人机系统搜索”，《智能机器人应用国际期刊》，第 2 卷，第 1 期，第 98–109 页，2018年。[26] G. Best、O. M. Cliff、T. Patten、R. R. Mettu 和 R. Fitch，“Dec-MCTS：面向多机器人主动感知的去中心化规划”，《国际机器人研究期刊》，第 38 卷，第 2–3 期，第 316–337 页，2019年。[27] P. Quin、D. D. K. Nguyen、T. L. Vu、A. Alempijevic 和 G. Paul，“在机器人探索中高效检测前沿单元的方法”，《机器人与人工智能前沿》，第 8 卷，第 616470 页，2021年。[28] M. T. J. Spaan、G. J. Gordon 和 N. Vlassis，“通信智能体团队在不确定性下的去中心化规划”，载于《第五届自主智能体与多智能体系统国际联合会议论文集——AAMAS ’06》，2006年。[29] T. H. Chung、G. A. Hollinger 和 V. Isler，“移动机器人学中的搜索与追逃：综述”，《自主机器人》，第 31 卷，第 4 期，第 299–316 页，2011年。[30] Y. Ma、R. Garnett 和 J. Schneider，“结合区域感知的稀疏信号主动搜索”，《AAAI 人工智能会议论文集》，第 31 卷，1，2017 no. 。[31] B. Jedynak、P. I. Frazier 和 R. Sznitman，“带噪声的二十个问题：熵损失下的贝叶斯最优策略”，《应用概率杂志》，第 49 卷，第 1 期，第 114–136 页，2012 年。[32] P. Rajan 、 W. Han 、 R. Sznitman 、 P. I. Frazier 和 B. M. Jedynak，“贝叶斯多目标定位”，载于《第 32 届国际机器学习会议论文集》第 37 卷，2015年，第 1945–1953 页。[33] Z.-Q. Zhao、P. Zheng、S.-T. Xu 和 X. Wu，“深度学习目标检测：综述”，arXiv [cs.CV], 2018。[34] R. E. Kalman，“线性滤波与预测问题的新方法”，《ASME 基础工程学报》，第 82 卷，D 系列，第 35–45 页，1960年。[35] 斯坦福人工智能实验室等，机器人操作系统。2018年。[36] M. R. O’Shaughnessy、M. A. Davenport 和 C. J. Rozell，“利用动态滤波进行稀疏贝叶斯学习，以推断时变稀疏信号”，《电气电子工程师学会信号处理汇刊》，第 68 卷，第 388–403 页，未定义 2020年。[37] P. Virtanen 等，“SciPy 1.0：用于科学计算的基础算法”，《自然方法》，第 17 卷，第 261–272 页，2020年。[38] C. Igoe、R. Ghods 和 J. Schneider，“多智能体主动搜索：一种强化学习方法”，《电气电子工程师学会机器人与自动化快报》，第 7 卷，第 2 期，第 754–761 页，2022年。[39] N. Koenig 和 A. Howard，“Gazebo 的设计与使用范式：一个开源多机器人仿真器”，2004年电气电子工程师学会/日本机器人学会智能机器人与系统国际会议（IROS）（电气电子工程师学会目录号 04CH37566），2004年，第 2149–2154 页，第 3 卷，doi: 10.1109/IROS.2004.1389727。[40]T. K. Moon，“期望最大化算法”，《电气电子工程师学会信号处理杂志》，第 13 卷，第 6 期，第 47–60 页，1996年，doi：10.1109/79.543975。