



# 无人机污染源自主搜寻算法综述

李中国, 陈文华\*, 刘存佳

Department of Aeronautical and Automotive Engineering, Loughborough University, Loughborough LE11 3TU, UK

\* 通信作者. E-mail: w.chen@lboro.ac.uk

收稿日期: 2022-01-27; 修回日期: 2022-02-21; 接受日期: 2022-03-18; 网络出版日期: 2022-09-05

英国工程和自然科学研究委员会 (EPSRC) (批准号: EP/T005734/1) 资助项目

**摘要** 自主搜寻在环境保护、灾后搜救和工业监控等方面有重要作用. 无人机传感平台凭借其灵活的机动性成为实现自主搜寻的有效途径之一. 针对自主搜寻中两个关键任务, 即污染源定位和污染边界跟踪, 本文分别梳理了相关算法体系, 并揭示了其设计思想的本质. 对于污染物溯源, 控制驱动和信息驱动两类算法是目前的研究主流, 但两者的结合却鲜有报道. 在边界跟踪方面, 本文从单平台和多平台协作两方面总结了现有文献的控制和估计方法. 根据双重控制思想, 本文分析了相关算法的设计思路, 并展望了下一步的研究重点.

**关键词** 自主搜寻, 无人机, 污染源搜寻, 边界跟踪, 双重控制, 探索与利用

## 1 引言

近年来, 多次重大灾难事故的发生深刻地凸显出污染源定位和污染边界监控的重要性, 具有代表性的案例包括 2010 年墨西哥湾漏油事故和 2011 年福岛第一核电站事故. 在污染物监测与定位方面, 早期文献主要聚焦在静态传感器网络的信息融合<sup>[1]</sup>. 特别是在具有较高风险的工业界, 例如石油天然气工厂、核电站等, 静态传感器网络被事先部署于潜在风险区域, 以便实现实时监控和及时干预. 此类方法的核心在于传感器网络的优化布局, 一旦部署完成将无法进行动态调整. 因此, 静态传感器网络主要适用于已知风险区域, 且部署成本较高<sup>[1,2]</sup>. 相反, 动态传感网络能够根据应用场景的不同, 灵活快速地部署到污染区域并执行相应任务.

自主搜寻 (autonomous search), 即利用可移动无人系统对未知环境中的未知目标进行定位和追踪. 近年来, 自主搜寻已被广泛地应用到工业生产、精准农业、环境监测、灾难管理、反恐搜救等相关领域<sup>[3,4]</sup>. 根据应用场景的不同, 常用的搜寻设备包括地面机器人 (unmanned ground vehicles, UGVs)、无人机 (unmanned aerial vehicles, UAVs) 以及水下机器人 (autonomous underwater vehicles, UAVs) 等.

**引用格式:** 李中国, 陈文华, 刘存佳. 无人机污染源自主搜寻算法综述. 中国科学: 信息科学, 2022, 52: 1579–1597, doi: 10.1360/SSI-2022-0044  
Li Z G, Chen W-H, Liu C J. Review of UAV-based autonomous search algorithms for hazardous sources (in Chinese). Sci Sin Inform, 2022, 52: 1579–1597, doi: 10.1360/SSI-2022-0044

由于机动能力和应用场景的限制,地面和水下机器人在自主搜寻应用方面受到一定的制约.随着无人飞行器技术的成熟,无人机设备成本得到有效降低,机动性和可靠性也逐步提升.因此,基于无人机构建的搜寻平台,如嗅探无人机(SNIFFDRONE<sup>[5]</sup>),不仅引起了空前的学术研究与讨论,也受到广泛的商业关注与投资.

无人机可以被用于广泛地搜寻场景,包括山地<sup>[6]</sup>、城区<sup>[7,8]</sup>、海洋<sup>[9]</sup>、火山<sup>[10,11]</sup>、丛林<sup>[12]</sup>等复杂环境<sup>[4,5]</sup>.除去无人机本身外,构建搜寻平台还需要传感器、数据处理器、控制决策等多个模块.实现自主搜寻需要涉及数据采集、传感器融合、环境感知和路径规划等多模块功能的协作.传感器的选择取决于待搜寻源头的属性,包括化学、生物、辐射及核物质(chemical, biological, radiological and nuclear materials, CBRN).一般来说,采集的数据往往包含大量噪声,甚至在极端情况下无法采集到有效的污染物浓度信息.为了应对数据噪声和不确定性所带来的棘手挑战,当前研究主要从两方面着手:硬件配置与算法设计.在硬件方面,受到无人机负载和续航方面的限制,通常搭载较为轻便简易的传感器.传感器安装过程中,除了要考虑复杂环境、污染物类型等一系列因素外,还要考虑到无人机自身所带来的局部湍流.文献<sup>[5]</sup>给出了详细的传感器性能及其布局综述.此外,增加传感器数量是另一类行之有效的方案.具体实施过程中,有单平台多传感器(multi-sensor single platform)和多平台多传感器(multi-sensor multi-platform)两类策略<sup>[13]</sup>.然而,复杂多模态传感器网络给数据融合与轨迹规划带来诸多前所未有的新挑战<sup>[13]</sup>.在算法方面,经典的搜寻策略包括化学趋向性(Chemotaxis)和风向趋向性(Anemotaxis),此类方法主要受生物行为的启发<sup>[14,15]</sup>.另一类典型的算法是从信息理论的角度出发的信息熵趋向性(Infotaxis),旨在收集更多环境和污染物信息,又称信息路径规划(informative path planning, IPP).近年来,随着计算能力和智能控制领域的突飞猛进,最近的文献中出现了一些全新的算法设计思路,典型的代表包括双重控制(dual control for exploration and exploitation, DCEE)和主动学习(active learning).

基于无人机的源搜寻平台的主要任务分为以下两类:源参数估计与溯源和边界跟踪.图1展示了本文结构和算法分类方法.源参数估计与溯源包含两个子任务,即通过传感器测量估计污染源的参数信息(包括位置和释放率等)和位置跟踪.源参数估计在许多应用中有至关重要的作用,具体讨论可参考文献<sup>[1]</sup>.另外,参数估计中的位置信息可为跟踪控制提供参考点.边界跟踪问题的目标略有不同,此类问题的目的在于追踪污染物扩散边界或者特定浓度的等位线.在污染物扩散中(如油气泄漏),跟踪边界可以用来估计污染范围、扩散速率,以便采取有效干预措施.边界追踪也在森林火灾、火山喷发等突发事件中扮演重要角色<sup>[3,4]</sup>.部署过程中一般涉及多无人系统的协同控制和动态资源调度.智能决策在各类任务中起到了至关重要的作用,可称之为核心,它决定了任务的成功与否、效率和性能.

值得注意的是,现有综述文章大多从无人机硬件平台和应用角度出发来归纳现有文献<sup>[3,5]</sup>.经典搜寻算法已被广泛收录在自主搜寻综述中,例如文献<sup>[14,16~18]</sup>.近年来,基于信息理论与控制理论的智能算法受到广泛关注,也在自主搜寻方面取得突破性进展.然而,最新的算法设计思想仍没有被包含在近年来的综述之中<sup>[13,14,16,17]</sup>.在门类繁多的应用中<sup>[3]</sup>,本文归纳出两大类搜寻任务(污染源定位和边界跟踪)以及相应算法.需要特别强调的是本文涵盖了双重控制和主动学习方面的工作,填补了现有综述文献在自主搜寻算法方面的缺失与不足.同时,基于该领域的最新突破,本文将提出自主搜寻方面的挑战、未来的研究方向和展望.

## 2 无人机平台与模型

自主搜寻的关键在于利用可移动无人机平台所采集到的浓度信息(通过化学/生物/放射性传感

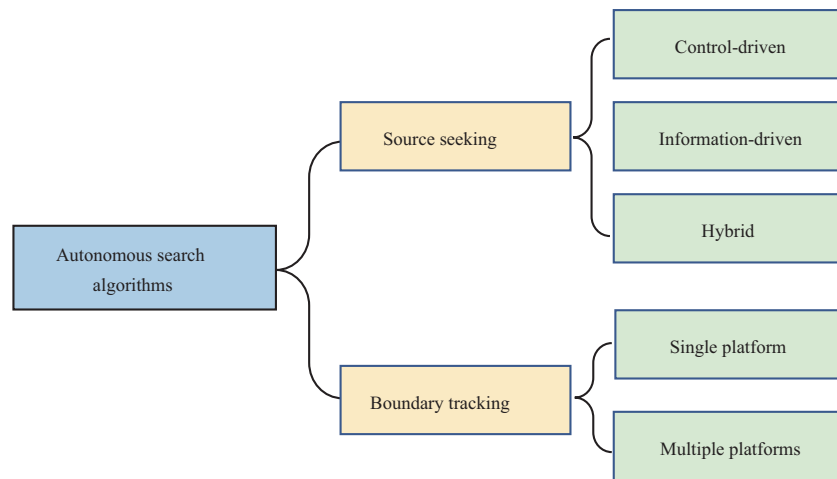


图 1 (网络版彩图) 自主搜寻任务与算法分类

Figure 1 (Color online) Classification of missions and algorithms for autonomous search

器), 规划无人机路径来实现在未知环境中对污染源进行定位追踪等功能. 这一过程涉及到无人机、传感器和未知环境之间的相互关联和作用.

## 2.1 无人机分类与建模

一般来说, 无人机搜寻任务采用分层控制和决策 (hierarchical control and decision-making), 主要特点是无人机底层控制和上层决策相互分离. 自主搜寻的相关研究主要聚焦在上层决策, 一般假设底层控制能够及时且准确地完成上层给定的跟踪目标. 因此, 可将无人机的模型简化为

$$\mathbf{p}_{k+1} = f(\mathbf{p}_k, \mathbf{u}_k), \quad (1)$$

其中,  $\mathbf{p}_k = [p_{k,x}, p_{k,y}, p_{k,z}]^T \in \mathbb{R}^3$  代表无人机在  $k$  时刻的位置矢量,  $\mathbf{u}_k \in \mathcal{U}_k \subset \mathbb{R}^3$  为  $k$  时刻的控制输入,  $\mathcal{U}_k$  定义为可行控制集,  $f(\mathbf{p}_k, \mathbf{u}_k)$  代表传递函数, 通常也可简化为  $\mathbf{p}_{k+1} = \mathbf{p}_k + \mathbf{u}_k$ .

考虑到本文的焦点不在无人机平台硬件方面, 所以对无人机类型仅作简要介绍. 综合经济成本、便捷性和安全性等多方面因素考虑, 用于自主搜寻任务的机体多为小型无人机, 一般不超过 25 kg (包括电池、传感器、控制设备等) [19]. 无人机一般分为旋翼无人机 (rotary-wing drones) 和固定翼无人机 (fixed-wing drones). 尽管固定翼无人机产生的额外湍流干扰较小, 但需维持高速飞行, 往往不能兼顾传感器的采样时间. 旋翼无人机会引入较严重的局部湍流, 这会严重影响传感器采集数据的准确性 (对于通过大气传播的化学污染物尤甚). 旋翼无人机可短暂停留在采样位置以增加数据的准确性, 所以在自主搜寻中较为常用 [5, 20]. 对于复杂时变气流和旋翼气流扰动下的无人机气味搜寻, 文献 [20] 给出了烟羽建模方法和实验验证.

## 2.2 污染物类型与扩散模型

常见的污染物扩散环境为大气和海水表面. 通过大气传播的应用范围更为广泛, 例如工业气体监控、搜救、火灾、火山灰及核污染物追踪等, 后者涉及的应用包括海洋油田泄漏监控、边界跟踪、海洋环境保护等. 大气污染物扩散的建模分析一直以来都是颇具挑战的问题. 扩散模型的选择需要根据精度要求、算力、颗粒物属性等多方面因素决定. 对于化学污染物, 常用的模型包括高斯烟羽模型 (Gaussian plume, GP), 计算流体动力学 (computational fluid dynamics, CFD) 等 [21]. 受局部气流

干扰影响, 大气中化学颗粒的扩散极为复杂, 因此呈现出高不确定性和难预测性. 各向同性扩散模型 (isotropic plume) 因其计算效率高又包含明确物理参数, 在诸多经典搜寻算法中得到广泛应用 [22~25]. 浓度分布  $\mathcal{M}(\mathbf{p}_k, \Theta_s)$  可写成

$$\begin{aligned} \mathcal{M}(\mathbf{p}_k, \Theta_s) = & \frac{q}{4\pi\zeta_{s1}\|\mathbf{p}_k - \mathbf{p}_s\|} \exp\left[\frac{-\|\mathbf{p}_k - \mathbf{p}_s\|}{\lambda}\right] \exp\left[\frac{-(p_{k,x} - s_x)u_s \cos\phi_s}{2\zeta_{s1}}\right] \\ & \times \exp\left[\frac{-(p_{k,y} - s_y)u_s \sin\phi_s}{2\zeta_{s1}}\right], \end{aligned} \quad (2)$$

其中,  $\Theta_s$  代表污染源和环境信息, 包括源位置矢量  $\mathbf{p}_s := [s_x, s_y, s_z]^T$ , 释放率  $q$ , 风速  $u_s$ , 风向  $\phi_s$ , 扩散率  $\zeta_{s1}$ , 颗粒寿命  $\zeta_{s2}$  及复合参数  $\lambda = \sqrt{\frac{\zeta_{s1}\zeta_{s2}}{1+(u_s^2\zeta_{s2})/(4\zeta_{s1})}}$ . 辐射物质的扩散较为不同, 辐射粒子受气流影响较小, 较为稳定. 因此, 辐射场的建模和追踪相对容易. 通常, 在单位时间内传感器所能检测到的辐射粒子与辐射源强度成正比, 与距离平方成反比 [26].

在自主搜寻方面也有一些算法不依赖于扩散模型, 典型的无模型算法包括极值搜寻 (extremum seeking), 此类算法采用梯度信息实时更新控制策略对目标函数进行优化. 事实上, 传统的化学趋向性算法也属于无模型算法, 它的本质是利用浓度梯度逼近最大含量位置. 此类无模型算法的鲁棒性与性能在复杂环境中的表现往往不尽如人意, 主要原因是梯度消失和梯度干扰问题. 不过近年来, 出现了一些研究尝试结合有模型和无模型算法, 如文献 [27, 28], 相关结果表明两者结合能够有效提升搜寻任务的成功率和性能.

### 2.3 传感器分类与建模

由于小型无人机载重、续航和成本限制, 机载传感器一般较为轻便简易, 受环境和自身噪声干扰明显. 根据污染物类型的不同, 相应的传感器也可分为化学、生物和辐射三类, 尤其化学和辐射传感器较为常用. 目前, 许多研究逐步增加数据采集的类型, 不仅包括浓度传感器, 也为无人机配备其他感知模块, 如风速风向传感器 [27, 29]、雷达 [30] 及图像传感器 [31~34] 等. 此类基于多模态多传感器的搜寻平台部署成本高昂、计算复杂、延展性和灵活性较差, 但在复杂场景中能够收集和利用更多环境信息, 文献 [13] 给出了较为完备的综述.

本文主要讨论基于浓度传感器的搜寻策略. 由于环境干扰、传感器噪声等因素影响, 浓度读数包含随机噪声. 在污染物浓度较低的情况下, 传感器可能无法检测到任何浓度信息. 文献 [5] 列举了一些常用的传感器类型及其性能. 在自主搜寻文献中 [23, 24, 28, 35, 36], 通常将传感器浓度表示为

$$z_k(\mathbf{p}_k) = \begin{cases} \mathcal{M}(\mathbf{p}_k, \Theta_s) + v_k, & D = 1, \\ \bar{v}_k, & D = 0, \end{cases} \quad (3)$$

其中,  $D = 1$  表示传感器采集到有效浓度信息, 而  $D = 0$  表示当前位置没有污染颗粒或者传感器未能检测出有效浓度信息,  $v_k, \bar{v}_k$  代表传感器噪声. 在实际应用中, 通常可根据传感器特性对事件  $D$  发生的概率  $p(D)$  做假设 [35]. 本文用向量  $\mathcal{Z}_k := \{z_1(\mathbf{p}_1), z_2(\mathbf{p}_2), \dots, z_k(\mathbf{p}_k)\}$  表示第一步到当前时刻  $k$  的历史测量数据. 在下文中, 除反应式算法 (reactive algorithms) 仅依赖当前测量做决策, 其他自适应或学习算法一般基于历史数据进行参数估计和决策.

## 3 源参数估计与跟踪 (source term estimation and tracking)

受双重控制思想的启发 [24], 本文将源参数估计和跟踪的方法大致分为三类: 以控制为导向的方

表 1 源参数估计与跟踪方法总结<sup>a)</sup>

Table 1 Summary of source term estimation and tracking methods

Type	Ref.	Algorithm	Estimated parameter	Estimation method	Decision-making method	Computational load
Control-driven	[37]	Chemotaxis	-	-	Gradient	Low
	[27]	Anemotaxis	$p$	Particle filter	Extensive search	High
	[24]	MPC	$p, q, u_s, \phi_s, \zeta$	Particle filter	Extensive search	High
	[6, 38]	ES	-	-	Estimated gradient	Low
	[39]	SPSA	-	-	Estimated gradient	Low
	[40]	Adaptive ES	Unknown parameters	Estimated gradient	Estimated gradient	Low
Information-driven	[22, 41]	Infotaxis	$p$	Bayesian	Extensive search/gradient	High
	[25]	Infotaxis II	$p, q, u_s, \phi_s, \zeta$	Particle filter	Extensive search	High
	[42]	FIM	$c, s$	Least-square	Nonlinear optimisation	Fair
	[23]	Entrotaxis	$p, q, u_s, \phi_s, \zeta$	Particle filter	Extensive search	High
	[2, 36]	KL divergence	$p, q, u_s, \phi_s, \zeta$	Particle filter	Extensive search	High
	[43]	Adaptive active planning	Unknown parameters	EKF	Nonlinear optimisation	Fair
Hybrid	[24, 44]	DCEE	$p, q, u_s, \phi_s, \zeta$	Particle filter	Extensive search	High
	[7]	Informed tree planning	$p, q, u_s, \phi_s, \zeta$	Particle filter	Extensive search	High
	[28]	CL-DCEE	$p, q, u_s, \phi_s, \zeta$	Concurrent least square	Gradient	Low
	[33, 45, 46]	Bayesian optimisation	Unknown parameters	Nonlinear interpolation	Nonlinear optimisation	High

a) SPSA: simultaneous perturbation stochastic approximation; FIM: fisher information matrix; KL divergence: Kullback-Leibler divergence; CL-DCEE: concurrent learning-based dual control for exploration and exploitation;  $c, s$  are concentration and source terms in the advection-diffusion model [42, 47].

法、信息路径规划, 及组合算法. 从该角度出发, 传统控制方法采用了纯利用策略 (exploitation), 信息路径规划则属于纯探索策略 (exploration), 而组合算法实现了利用与探索的平衡 (balance between exploitation and exploration). 表 1 [2, 6, 7, 22~24, 24, 25, 27, 28, 33, 36~42, 42~47] 给出了参数估计与跟踪方法的总结, 具体分类依据将在下文中详细阐述.

事实上, 现有文献大多将源参数估计和跟踪问题区分开来. 前者指的是通过路径规划和浓度测量重构污染源参数信息 (如释放速率、位置信息等); 后者指的是规划路径使得无人机接近污染源位置并进行跟踪<sup>[1]</sup>. 从本质上说, 重构污染源信息是参数估计的过程, 而污染源跟踪属于控制问题. 相应地, 这导致了两种独立的决策思想: 信息驱动策略和控制驱动策略. 也正是这种分类方式使得现有搜寻算法或偏向于纯利用、或偏向于纯探索. 图 2 整合了双重目标驱动的搜寻任务, 决策过程需同时考虑参数估计和追踪两个子任务的性能指标. 无人机的下一步位置直接决定了将要采集到的数据, 而这些含有噪声的数据 (位置与浓度) 又将影响源参数估计和跟踪的性能指标. 下面引入两个子任务的定义.

**定义 1 (源参数估计)** 对于给定历史测量  $\mathcal{Z}_k := \{z_1(\mathbf{p}_1), z_2(\mathbf{p}_2), \dots, z_k(\mathbf{p}_k)\}$ , 源参数估计旨在利用扩散模型和先验知识重构污染源信息  $\Theta_s$ , 一般描述为概率分布函数  $p(\Theta_k | \mathcal{Z}_k)$ .

**定义 2 (轨迹优化)** 给定目标函数  $J(\cdot)$  和当前概率分布函数  $p(\Theta_k | \mathcal{Z}_k)$ , 最优路径规划旨在选择可行控制量  $\mathbf{u}_k^* \in \mathcal{U}_k$  使得目标函数最小, 即  $\mathbf{u}_k^* = \operatorname{argmin}_{\mathbf{u}_k \in \mathcal{U}_k} J(\mathbf{u}_k)$ .

虽然可将源参数估计和跟踪问题看作自适应控制, 但该类问题与传统自适应控制有一点明显区别: 传统自适应控制一般估计系统未知参数或状态<sup>[48, 49]</sup>, 而源参数估计问题实际上是对未知环境的估计. 污染源是未知环境的一部分, 独立于无人机系统之外. 事实上, 通过学习污染源位置为无人机提供了待追踪的参考轨迹, 而在传统自适应控制中参考轨迹一般预先给定.

### 3.1 控制驱动策略

控制策略一般构建物理意义明确的性能指标, 如跟踪误差, 并通过优化该指标驱动无人机完成相应的任务. 因此, 在早期文献中这类算法也被称为任务驱动型策略 (task-driven methods)<sup>[50]</sup>. 总的来说, 控制相关的算法既包含了传统的反应式策略 (reactive strategy), 也涵盖了自适应算法 (adaptive strategy). 前者的经典例子有 Chemotaxis<sup>[37]</sup> 和 Anemotaxis<sup>[27, 51]</sup>; 而自适应类控制算法包括模型预

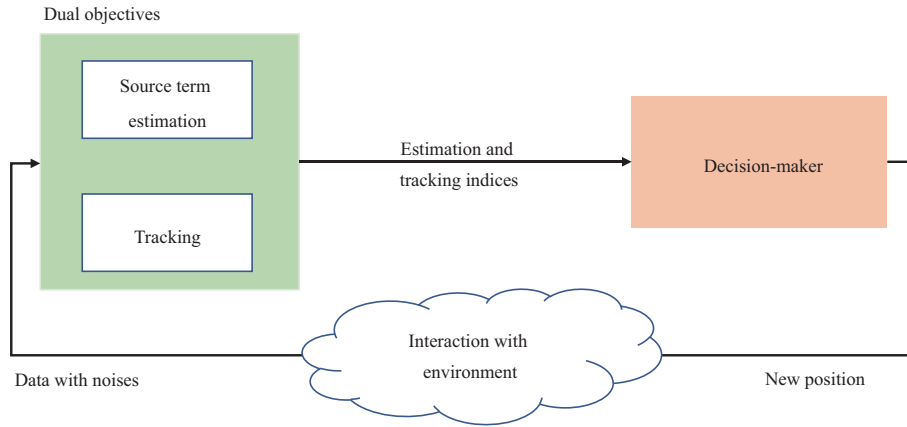


图 2 (网络版彩图) 双重目标决策过程: 源参数估计与跟踪

Figure 2 (Color online) Dual-objective-driven decision-making: source term estimation and tracking

测控制 (model predictive control, MPC) [24, 52, 53]、自适应极值搜索 (extremum seeking, ES) [43] 等. 需要特别指出的是, 反应式算法仅利用即时测量数据来驱使无人机靠近污染源位置, 该类策略往往不进行源参数估计. 当然, 也有一些研究通过历史数据学习污染源信息, 例如文献 [27] 将贝叶斯 (Bayes) 估计与风趋向算法结合. 本文主要回顾基于自适应和参数估计的控制算法, 此类算法的参数估计过程实际上是被动学习 (passive learning). 换言之, 无人机决策过程中只考虑跟踪指标 (参考图 2), 但未来的数据会间接地影响参数估计, 这也是传统自适应控制和系统辨识所采用的方法 [48, 54].

模型预测控制 (MPC) 虽已被广泛应用到无人机控制中, 但相关文献大多关注轨迹跟踪和调节问题 [53, 55, 56], 本质上属于底层控制. MPC 能够灵活处理各类系统和环境约束 (如输入输出限制、避障等), 且可保证对既定指标的优化, 因此可用于复杂环境下的自主搜寻. 根据文献 [24], 可将自主搜寻任务描述为

$$\begin{aligned} \min_{\mathbf{u}_k \in \mathcal{U}_k} J(\mathbf{u}_k) &= \mathbb{E}_{\Theta_s} \left[ \|\mathbf{p}_{k+1|k} - \mathbf{p}_s\|^2 \mid \mathcal{Z}_k \right] \\ \text{s.t. } \mathbf{p}_{k+1|k} &= f(\mathbf{p}_k, \mathbf{u}_k), \end{aligned} \quad (4)$$

其中,  $\mathcal{U}_k$  表示可行控制集, 这里控制决策既需要考虑无人机自身安全限制 (例如最大速度限制), 也需要考虑环境限制与不确定性 (如障碍物、气流干扰等). 式 (4) 的物理解释为: 决策系统基于历史测量数据  $\mathcal{Z}_k$  给出对污染源位置的估计, 并从可行控制集中选取下一步控制量以减小无人机与估计位置的距离. 模型预测控制的本质在于利用现有信息选取最佳决策来实现控制目标. 在促使无人机向估计源位置移动的过程中, 所采集的浓度信息会间接地用于更新参数估计. 因此, MPC 属于纯利用 (exploitation) 策略. 文献 [24] 仅给出采用单步 MPC 处理自主搜寻问题的框架和实验结果, 未来研究方向可考虑如何完善该框架, 例如, 通过多步 MPC 结合其他参数估计方法改进搜寻性能、降低计算量等.

极值搜索是经典的无模型优化, 用来处理输入 - 输出模型未知的控制问题 [57, 58]. 近年来, 极值搜索被逐渐应用到诸多工程问题中, 如刹车系统、工业过程控制、发动机控制等 [59]. 极值搜索的思想也被引进到自主搜寻任务中 [6, 60]. 究其根本, 自主搜寻可以理解为将无人机控制到未知浓度分布的最大值所在位置. 随着极值搜索理论框架的逐步完善, 基于该方法的自主搜寻研究一般都给出了完备的收敛性、稳定性分析 [38, 60]. 这与基于信息驱动类策略形成鲜明对比, 由于缺乏理论分析工具, IPP 的理论保障仍然是亟待解决的问题, 后文中将详细阐述. 文献 [39] 提出了基于随机梯度估计的搜寻方法, 并分别对有障碍物和无障碍物两种环境下的算法给出了完备的理论分析. 文献 [6] 采用极值搜索的方

法控制无人机位置, 用来实现雪崩事件中的救援问题. 该研究假设受困人员配备有可发射无线电波的紧急救援装置, 无人机通过机载信号接收装置进行自主路径规划以尽快找出被困人员位置. 无人机协同控制的飞速发展也推动了多平台协作搜寻的进步. 多平台编队控制能够有效突破单平台的局限、促进污染源快速定位<sup>[51,61~65]</sup>. 多无人机平台在多污染源同时跟踪问题中是必不可少的, 其协作方式可分为集中式和分布式两类<sup>[66]</sup>. 相应的优缺点将在下文第 4 节中具体展开.

自适应极值搜索<sup>[40]</sup>考虑被控系统存在输入-输出参数化模型, 通过参数估计重构该模型并利用被估参数设计极值搜索算法. 在自适应极值搜索中存在两个目标, 即参数估计和极值搜索, 这实际上构成了一种有模型(参数未知)的极值搜索. 为保证参数估计收敛, 文献[40]给出了持续激励的条件, 但控制决策仅考虑输出最大化, 并未将参数估计的性能考虑到决策过程中. 由于参数估计与极值搜索控制器均采用梯度下降的算法<sup>[40]</sup>, 计算复杂度适中, 但现有研究尚未将此类自适应算法引入到自主搜寻中. 极值搜索中引入的摄动(正弦或随机信号)旨在估计梯度信息而非探索环境, 因此这些算法本质上仍属于纯利用的控制策略.

总体来说, 控制驱动策略目标在于将无人机移动到污染源位置. 实现这一控制目标既可通过启发式梯度下降算法(如 Chemotaxis 和极值搜索)亦可通过估计污染源位置为无人机提供跟踪参考(如自适应极值搜索和 MPC). 两者都体现了对已知信息的完全利用: 前者利用当前梯度测量或梯度估计(存在噪声), 后者利用现有污染源的位置估计(存在不确定性). 受噪声与不确定性影响, 纯利用(exploitation)搜寻算法的性能与鲁棒性广受质疑<sup>[16,17,25]</sup>.

### 3.2 信息驱动策略

信息驱动策略的本质在于选择最优决策实现对特定信息测度的最优化. 常用的信息测度包括方差、信息熵、KL 散度(Kullback-Leibler divergence)等<sup>[67]</sup>. 由于下一步控制所带来测量的不同, 它对当前认知(current belief)的影响也不尽相同. 信息测度最优化的过程本质上是探索环境的过程, 即减少环境信息的不确定性.

在信息驱动算法中, 不少研究工作采用参数化滤波器<sup>[43,68]</sup>(如卡尔曼滤波、EKF、UKF 等), 此类算法的主要优势有两点: 其一是计算负荷相对较低, 其二是算法迭代过程直接提供了信息不确定性测度(如卡尔曼滤波器中的协方差矩阵)<sup>[13]</sup>. 参数化滤波器的性能依赖于模型的准确性、分布密度的高斯性和测量值的非稀疏性. 然而, 真实的搜寻任务并不满足线性和高斯条件, 通常情况下浓度测量有稀疏的特点, 使得复杂环境中的参数估计更加棘手. 因此, 目前的趋势倾向于采用贝叶斯估计框架. 在此基础上, 源参数估计在时刻  $k$  的后验概率为  $p(\Theta_k | \mathcal{Z}_k)$  (定义 1). 给定下一个采样位置及其浓度测量  $z_{k+1}(\mathbf{p}_{k+1})$ , 参数估计的后验概率为

$$p(\Theta_{k+1} | \mathcal{Z}_{k+1}) = \frac{p(z_{k+1} | \Theta_{k+1})p(\Theta_{k+1} | \mathcal{Z}_k)}{p(z_{k+1} | \mathcal{Z}_k)}, \quad (5)$$

其中  $p(z_{k+1} | \mathcal{Z}_k) = \int p(z_{k+1} | \Theta_{k+1})p(\Theta_{k+1} | \mathcal{Z}_k)d\Theta_{k+1}$ ,  $\mathcal{Z}_{k+1} := \{z_1(\mathbf{p}_1), z_2(\mathbf{p}_2), \dots, z_k(\mathbf{p}_k), z_{k+1}(\mathbf{p}_{k+1})\}$ . 基于  $k$  和  $k+1$  两时刻的后验概率, 新的浓度测量  $z_{k+1}(\mathbf{p}_{k+1})$  所带来的信息增益可定义为信息熵的差, 即

$$\mathcal{I}_{\text{gain}}(z_{k+1}(\mathbf{p}_{k+1})) := \mathcal{H}(p(\Theta_k | \mathcal{Z}_k)) - \mathcal{H}(p(\Theta_{k+1} | \mathcal{Z}_{k+1})), \quad (6)$$

其中,  $\mathcal{H}(p(\Theta_k | \mathcal{Z}_k))$  和  $\mathcal{H}(p(\Theta_{k+1} | \mathcal{Z}_{k+1}))$  分别为采样  $z_{k+1}(\mathbf{p}_{k+1})$  前后的信息熵<sup>[67]</sup>. 熵是一种描述不确定性较为常用的测度, 对于给定概率密度函数  $p(\Theta_k | \mathcal{Z}_k)$ , 其定义为  $\mathcal{H}(p(\Theta_k | \mathcal{Z}_k)) =$



$-\int p(\Theta_k | \mathcal{Z}_k) \ln(p(\Theta_k | \mathcal{Z}_k)) d\Theta_k$ . 对于高斯分布, 熵和方差  $\sigma^2$  满足  $\mathcal{H} = \frac{1}{2} (1 + \ln(2\pi\sigma^2))$ . 因此, 假使未知参数估计  $p(\Theta_k | \mathcal{Z}_k)$  服从高斯分布, 采用熵或方差的路径规划算法等价<sup>[69]</sup>.

基于信息熵的算法包括经典的 Infotaxis<sup>[22, 25, 41]</sup> 和 Entrotaxis<sup>[23, 70]</sup>. 事实上, Infotaxis 遵从信息增益定义 (6), 即选取能够最大化信息增益的采样点, 而 Entrotaxis 的本质是将无人机移动到信息最不确定的位置, 根据预测熵值

$$\mathcal{I}_{\text{entropy}} = - \int p(z_{k+1}(\mathbf{p}_{k+1}) | \mathcal{Z}_k) \ln(p(z_{k+1}(\mathbf{p}_{k+1}) | \mathcal{Z}_k)) dz_{k+1}. \quad (7)$$

Vergassola 等<sup>[22]</sup> 意识到在污染源定位中, 浓度测量往往具有较高的稀疏性. 传统基于梯度的搜寻方法 (如 Chemotaxis) 难以实现鲁棒的性能. 基于贝叶斯估计和信息测度的搜寻方法能够在测量稀疏的情况下实现优越的实验结果<sup>[71]</sup>. 在此之后, 许多基于 Infotaxis 的改进算法应运而生, 例如文献<sup>[22, 25, 41]</sup>. Hutchinson 等<sup>[23]</sup> 认为探索 (exploration) 的本质应是对未知区域的开拓, 也就是不确定性采样 (uncertainty sampling), 并基于这一思想提出了全新的 Entrotaxis 算法.

KL 散度可用来衡量两个分布密度的差异 (difference), 也被称为相对熵<sup>[67]</sup>, 定义为

$$\mathcal{I}_{kl}(p(\Theta | \mathcal{Z}_{k+1}) \| p(\Theta | \mathcal{Z}_k)) = \int p(\Theta | \mathcal{Z}_{k+1}) \ln\left(\frac{p(\Theta | \mathcal{Z}_{k+1})}{p(\Theta | \mathcal{Z}_k)}\right) d\Theta. \quad (8)$$

KL 散度涉及采样前后参数估计的后验概率分布密度, 根据上述定义可知该测度具有不对称性, 即  $\mathcal{I}_{kl}(p(\Theta | \mathcal{Z}_{k+1}) \| p(\Theta | \mathcal{Z}_k)) \neq \mathcal{I}_{kl}(p(\Theta | \mathcal{Z}_k) \| p(\Theta | \mathcal{Z}_{k+1}))$ . 但这两种信息定义均可用来构建决策目标函数 (采样前后的相对信息熵), 其物理含义有一定模糊性. Hutchinson 等<sup>[2, 36]</sup> 将基于 KL 散度的信息路径规划应用到源参数估计和搜寻任务中, 相关实验结果也展示出无人机平台在工业监控中的无限潜力.

信息驱动算法具有较强的鲁棒性, 适用于强非线性、非高斯和稀疏估计, 目前在自主搜寻方面已成为主流. 然而, IPP 依赖粒子滤波算法且轨迹规划和参数估计深度耦合<sup>[2, 22, 23, 25]</sup>, 这造成了 IPP 算法目前只进行单步规划 (one-step ahead), 并且可行控制集较小. 虽然近期的研究带来一些路径规划上的创新<sup>[7]</sup>, 但巨大的计算负荷逐渐成为亟待突破的发展瓶颈. 当前, 多数实验设备仍然需要远程计算中心协助无人机平台进行轨迹规划和贝叶斯估计<sup>[2, 35, 36, 44]</sup>, 这限制了无人机平台的灵活性和普适性.

信息测度的定义均基于参数估计的后验概率分布, 相关算法旨在从环境中汲取更多信息用来重构环境参数. 然而, 最具信息的位置 (或不确定性最大的位置) 并不一定是污染源所在的位置. 尽管信息驱动策略能够实现优越的环境估计, 但这一本质差别导致此类算法难以获得满意的跟踪效果<sup>[24]</sup>. 另外, 虽然已有许多文献尝试解释诸多信息测度之间的关联和物理含义, 但大多针对特定分布及特定信息函数<sup>[67, 72]</sup>.

### 3.3 组合平衡算法

通过上述对控制驱动和信息驱动两类算法的阐释, 我们注意到控制目标 (跟踪) 和信息目标 (参数估计) 既不对立、也不统一. 具体来说, 控制目标在于减小无人机与污染源的距離, 信息目标在于探索未知环境减少估计不确定性. 过度偏向于利用 (exploitation) 易造成参数估计不准确、跟踪误差较大或无法找到污染源; 而偏向于探索 (exploration) 易造成时间成本、控制成本的浪费, 且跟踪性能较差.

在机器学习研究中出现了大量工作尝试构建统一框架来实现探索与利用的平衡<sup>[73~78]</sup>. 强化学习 (reinforcement learning, RL) 所考虑的问题正是系统与环境的交互, 即无人系统采取行动并观测该行动在环境中的奖惩. 强化学习采用大量反复试验训练深度神经网络对状态-行动-奖惩 (state-action-reward) 进行遍历性学习, 由此建立完备的全局认知网络. 目前强化学习的主要应用仅限于静态环境



下、可进行大规模重复尝试的应用,如 Alpha-Zero<sup>[79]</sup>.在无人机系统或其他机械系统中,大规模重复尝试一般不可行:其一重复实验耗时、耗材;其二某些全局搜索往往违背物理系统的安全性、稳定性限制.针对自主搜寻任务,应用场景、污染物类型、风速风向、传感器类型等一系列因素都会呈指数级扩展状态-行动-奖惩的维度,其复杂度远远大于 Alpha-Zero 的有限状态空间.这也造成了强化学习通过实验所训练的模型难以应用到真实搜寻场景中,即迁移性较差.文献[80]详细阐释了强化学习在控制应用中的局限性.

在控制理论中,关于平衡探索与利用的历史可追溯到 1960 年, Feldbaum<sup>[81~84]</sup> 认识到控制输入不仅能够直接影响系统状态,也会影响未知参数的不确定性,因此称其为双重控制 (dual control). Bar-Shalom 和 Tse<sup>[85]</sup> 给出了双重控制的定义,并探讨了主动学习与被动学习的关系与不同.所谓双重控制指的是控制输入不仅能够影响系统状态,同时可以影响系统状态的  $r$  阶中心距 ( $r > 1$ ).区别于传统反馈控制,双重控制将未来的观测值应用到当前时刻的控制设计中,因此形成闭环动态规划问题.然而,至今复杂动态规划问题仍没有较好的解决方案<sup>[86~88]</sup>.尽管双重控制的概念新颖,但其计算复杂度较高,因此并未受到足够关注.近年来,随着计算能力的提升以及学习算法的突破,双重控制也重新受到广泛关注,典型的代表包括基于主动学习的模型预测控制<sup>[89]</sup>和强化学习<sup>[73]</sup>.

Chen 等<sup>[24]</sup>创新地将双重控制 (DCEE) 的概念引入到自主搜寻任务中,并提出了以贝叶斯估计为框架的解决方案.该方法的数学描述为

$$\begin{aligned} \min_{\mathbf{u}_k \in \mathcal{U}_k} J(\mathbf{u}_k) &= \mathbb{E}_{\Theta_s} \left[ \|\mathbf{p}_{k+1|k} - \mathbf{p}_s\|^2 \mid \mathcal{Z}_{k+1|k} \right] \\ \text{s.t. } \mathbf{p}_{k+1|k} &= f(\mathbf{p}_k, \mathbf{u}_k). \end{aligned} \quad (9)$$

对比式 (4) 中 MPC 的描述,其关键不同在于 DCEE 将下一步预测浓度引入到环境估计中,意味着当前的决策  $\mathbf{u}_k$  将直接影响参数估计的效果.通过大量对比仿真与实验,双重控制显示出优越的搜寻性能.需要特别指出的是,最初双重控制的概念仅关注系统状态本身,而忽略了环境因素<sup>[81~85]</sup>.Chen 等<sup>[24]</sup>将双重控制的意义进行了扩展,将其推广到系统与环境的关联.由于污染源参数本身属于环境信息,实现对环境信息的感知实际上为无人机提供了目标参考点.这与传统控制有重大区别:传统控制系统一般可分为跟踪和调节问题,相应的参考工作点已经事先给定,所有的环境信息均被视为系统干扰.而在自主搜寻中<sup>[24]</sup>,无人系统需实时感知环境、估计跟踪目标,本质上形成了一种先进的目标驱动控制 (goal-oriented control).所谓目标驱动指的是自主系统不依赖于人为干预、不依赖于既定参考轨迹,相反地能够完全自主地实现控制目标 (如估计污染源信息并跟踪).该方法的里程碑意义不仅在于实现了控制驱动与信息驱动两类算法的统一,也展示出了双重控制卓越的实践意义,这将推动双重控制理论在其他领域的快速发展.目前,基于这一理论框架, Rhodes 等<sup>[44]</sup>已将双重控制应用到无人机在复杂环境中的源参数估计和跟踪.文献[7]受双重控制和基于采样路径规划算法的启发,解决了复杂扩散、城市环境下的自主搜寻.目前 DCEE 均基于贝叶斯估计框架 (粒子滤波),对算力要求较高,且缺少理论分析工具.因此,文献[28]提出了基于多估计器同步学习的双重控制算法,并利用优化与控制工具给出了初步的稳定性分析,对未来建立更为完备的理论框架提供了参考.这一算法继承了 DCEE<sup>[24]</sup>粒子滤波的鲁棒性,又吸收了梯度下降的高效性,对于扩展单步到多步的双重控制有重要意义.

事实上,在自适应控制领域已有一些初步结果尝试平衡探索与利用.在不同文献中,所用术语也不尽相同,最为常见的是主动学习 (active learning) 和主动感知 (active sensing),旨在将参数估计、状态估计或环境感知的性能引入到决策指标中.在许多文献中这样的探索信号常常被刻意地加入到控制指标中<sup>[89,90]</sup>,此类做法又为控制设计引入了额外的超参选择,使得控制器设计更加困难.这也体现

了 DCEE 作为统一算法的另一个优点: 它避免了人为引入平衡系数, 该方法所导出的性能指标是具有明确物理意义的最优控制<sup>[24]</sup>. 被动学习本质在于利用现有信息 (certainty equivalence) 选取最佳决策来实现控制目标. 在促使无人机向估计源位置移动的过程中, 所采集的浓度信息会间接地用于更新参数估计. 主动学习是指无人机在决策过程中考虑下一步移动对参数估计的影响, 直接减少环境信息的不确定性. 一般来说, 强化学习与自适应 MPC 采用被动学习的方式更新未知参数. 双重控制<sup>[24]</sup>中融入了探索 (信息目标) 和跟踪指标 (控制目标), 因此属于主动学习. 相比于强化学习, DCEE<sup>[24]</sup>不需要进行大规模循环尝试, 无人机搜寻过程中仅聚焦于当前位置下的控制决策, 因此计算负荷低, 搜寻效率高.

虽然一些学者认为可将贝叶斯优化 (Bayesian optimisation) 用于自主搜寻<sup>[33, 46]</sup>或主动感知<sup>[13]</sup>, 但截至目前还没有充分证据表明此类方法的有效性, 我们对其适用性也有一定保留. 一般来说, 贝叶斯优化通过构建代理模型 (surrogate model) 对黑箱函数进行建模并优化. 鉴于黑箱模型计算复杂度高, 贝叶斯优化希望通过少量采样构建足够准确的代理模型. 这一过程需定义额外的获得函数 (acquisition function) 并通过优化该函数选取下一采样点, 获得函数的选取决定了能否实现利用与探索的平衡. 对于自主搜寻来说, 这样的做法使得超参选择困难、问题更加复杂. 另外, 贝叶斯优化作为黑箱优化算法, 完全忽视污染物扩散模型、环境模型和已知信息等, 其搜寻效率有待考量<sup>[45]</sup>. 更为关键的问题在于获得函数的定义是针对整个决策空间的, 对该函数施加系统模型限制 (即  $p_{k+1|k} = f(p_k, u_k)$ ) 或将严重影响其性能.

综上所述, 因控制驱动类算法易造成无人机对环境认知不足或存在较大偏差, 其搜寻效果在环境不确定性或干扰较大时不尽如人意<sup>[16]</sup>. 经过大量实验验证, 信息驱动类算法能够实现较为准确的污染源参数估计, 算法鲁棒性强, 搜寻成功率高<sup>[2, 22, 23, 35, 36]</sup>. 由于信息驱动类算法过度强调信息增益, 而信息增益最大的位置与污染源所在位置并无直接关系, 故其跟踪误差往往较大. 此外, 过度倾向于探索易造成控制成本浪费、搜寻时间提高. 平衡两类算法的关键在于对探索与利用的审慎权衡, 有效的平衡能够提升环境估计的准确性、搜寻成功率、算法鲁棒性并减少时间成本.

#### 4 边界跟踪 (boundary tracking)

边界跟踪指的是无人系统在污染边界或特定等位线进行巡航跟踪. 在环境保护方面有许多重要作用, 例如监控油气扩散、估计森林火灾范围、跟踪火山灰分布以及核污染区域等. 为了实现较好的边界跟踪, 数据采集往往需具备一定的空间分布特性, 例如环绕污染边界. 这使得单一平台的搜寻效果大打折扣, 在处理大面积污染扩散事件中, 单一平台更为捉襟见肘. 不同于静态单一污染源定位, 多无人机协同监控在边界跟踪或广义覆盖控制任务中更为常见. 因此, 下文按照平台数量分为单平台和多平台两类. 相比于传统的中心化协调方案, 分布式控制 (distributed control) 具有较强的扩展性、灵活性和鲁棒性等优点, 在无人机协作边界跟踪中已崭露头角. 具有代表性的控制算法包括 Bang-bang 控制、滑模控制 (sliding mode control, SMC) 和 MPC. 文献 [91] 对于复杂环境下的路径规划问题给出了较为完备的总结, 并着重讨论了无人系统避障与协同控制的难点与解决方案. 目前, 许多研究的决策过程需通过历史测量数据对边界范围进行估计和预测, 然后操控无人平台跟踪估计边界. 边界估计的算法多种多样, 较为常用的有累计和滤波器 (cumulative sum filter, CUSUM)、支持向量机 (support vector machine, SVM), Kringing 插值等学习与优化方法. 表 2<sup>[92~110]</sup> 列出了具有代表性的方法总结, 详细介绍将第 4.1 和 4.2 小节中展开.

表 2 边界跟踪方法总结

Table 2 Summary of boundary tracking methods<sup>a)</sup>

Platform	Ref.	Control algorithm	Boundary estimation	Boundary type	Source	Structure	Computational load
Single	[92]	Bang-bang	Grid-based RBE	Static	Radiation	-	Fair
	[93]	Bang-bang	Spline interpolation	Dynamic	Atmospheric dispersion	-	Low
	[94]	SMC	-	Static	bathymetry isoline	-	Low
	[95]	SMC	-	Dynamic	Volcanic ash	-	Low
	[96, 97]	SMC	-	Static	Scalar field	-	Low
	[98]	MPC	Universal Kriging	Dynamic	Oil spill	-	High
	[99]	Gradient-descent	Observer	Dynamic	Airborne plume	-	Low
	[100]	Proportional control	Uniform interpolation	Static/slowly moving	Environmental	-	Low
	[101]	Robust control	RBF NN	Static	Environmental	-	High
	Multiple	[102]	Bang-bang	-	Static	Scalar field	Centralized
[103~105]		Bang-bang	CUSUM	Static	Scalar field	Centralized	Fair
[106, 107]		Proportional control	GSO	Static	Environmental	Centralized	High
[108]		Formation control	Kalman filter	Static	Scalar field	Centralized	Fair
[109]		Formation control	State observer	Dynamic	Airborne contaminant	Distributed	Fair
[110]		MPC	Correlation map	Static	Airborne contaminant	Distributed	High

a) RBE: recursive Bayesian estimation; RBF NN: radial basis function neural network.

#### 4.1 单平台边界跟踪

Bang-bang 控制 (或称启停式控制) 结构简单、意义清晰, 又符合污染区域内外的不同特点, 因此在早期文献中被广泛采用<sup>[102, 104]</sup>. 其基本控制方案可归纳为: 如果无人机处于边界内则向外移动, 反之向内. Kemp 等<sup>[102]</sup> 采用了即时浓度测量作为决策的依据, 并根据 Bang-bang 控制进行轨迹规划. 由于测量噪声和不确定性干扰, 即时测量往往不能精确反映当前时刻的真实浓度值, 这导致相关算法鲁棒性较差. 为解决这一问题, Jin 和 Bertozzi<sup>[103]</sup> 提出对含有噪声的历史测量数据进行累计和滤波, 以此估计边界位置. 该策略的有效性依赖于大量数据, 因此时效性较差. 并且该算法仅适用于静态边界 (文献 [103] 假设该边界为直线), 对动态边界估计效果较差. Brink<sup>[93]</sup> 对边界估计算法进行了改进, 提出了基于样条插值 (spline interpolation) 的边界估计方法, 并将其与 Bang-bang 控制结合用于动态边界跟踪问题. 类似的边界估计方法还有多边形均分插值 (uniformly polygon interpolation)<sup>[100]</sup>, RBF 神经网络<sup>[101]</sup>, 集群优化算法 (例如, glowworm swarm optimisation, GSO)<sup>[106, 107]</sup> 等.

滑模控制通过定义滑动曲面 (sliding surface), 配合大增益设计, 强迫动态系统沿着滑动曲面运动. 滑模控制与启停式控制类似, 控制设计简单, 无需精确, 且有较强的鲁棒性. 在边界跟踪问题中, 需根据待跟踪的浓度线定义相应的滑动曲面, 即  $S = \{\mathbf{p}_r \in \mathbb{R}^3 \mid \sigma(t) = 0\}$ , 其中,  $\sigma(t)$  称为滑动变量, 定义为测量浓度与参考浓度差值, 即  $\sigma(t) = z(\mathbf{p}_k) - b$ . 虽然近年来滑模控制器形式略有变化, 但整体设计思想大同小异, 例如  $\mathbf{u}_k = -\alpha \text{sign}(\sigma(t))$ . 该切换控制算法配合大增益设计易造成切跳效应 (chattering)<sup>[96]</sup>, Dong 和 You<sup>[97]</sup> 提出了类 PI 的滑模控制器并给出了完备的全局收敛证明, 在跟踪等位线的过程中能够避免一般滑模控制中的切跳震颤效应. Mellucci 等<sup>[94]</sup> 提出了次优 SMC 算法用以测量静态海水深度等位线, 并通过实验验证了其跟踪有效性. 在污染边界跟踪方面, Menon 等<sup>[95]</sup> 采用了滑模控制跟踪环境污染边界, 相关结果表明该算法可用于静态或缓慢移动的边界跟踪.

模型预测控制能够有效处理各类系统限制和环境限制, 在现代复杂控制系统中得到广泛应用<sup>[111]</sup>. MPC 在源参数估计和跟踪方面的进展及其优势已在前文中详细阐述. 实现动态边界跟踪同时保证系统安全 (如避障功能) 一直是颇具挑战的问题. 目前, 已有一些初步成果展示出模型预测控制的巨大潜力<sup>[98]</sup>. 然而, 模型预测控制对算力要求较高, 且依赖模型准确性, 对于复杂气体流体扩散的建模往往不能满足 MPC 的需求. 文献 [98] 采用了 Kriging 空间插值的方法对浓度分步进行重构, 该方法对于静态分布拟合效果较好. 核辐射有较为稳定的辐射场, 文献 [92] 采用了等位线跟踪的方法对辐射源进

行估计, 该方法将搜寻区域网格化并利用贝叶斯迭代的方法实现定位. 文献 [1] 回顾了其他相关边界估计算法, 例如神经网络、支持向量机和进化算法等. 综合现有文献的研究趋势来说, 估计与控制相结合的边界跟踪算法是目前的研究主流, Jiang 和 Li<sup>[99]</sup> 梳理了在污染边界搜寻的主要设计框架, 并在此基础上提出了估计和控制的统一框架, 称为交互估计控制 (interactive estimation and control). 该框架采用了梯度下降的观测器和控制器, 虽然观测器被引入控制回路, 但其本质思想属于被动估计 (passive estimation).

## 4.2 多平台协作边界跟踪

相比单平台跟踪, 多无人机协作至少有以下四方面优势: (1) 对污染区域的有效覆盖; (2) 同步数据采集更具时空多样性; (3) 时效性强, 在大范围跟踪任务中减少了无人机远距离调度; (4) 鲁棒性强, 多平台测量对噪声和干扰抑制效果提升, 有利于精准边界估计. 尽管如此, 多无人机平台也不可避免地带来诸多新挑战: (1) 多无人机调度更复杂, 需考虑边界覆盖、防碰撞、相互协调、容错等因素; (2) 无人机间信息交互困难, 极端环境下的通信渠道、带宽限制较强; (3) 计算复杂度大幅增加, 多无人机协作过程中不仅要处理大量时空分布的浓度测量, 也要处理无人机间交互、规划协调信息. 目前, 可将多无人机平台协作分为集中式和分布式两种通信方式. 所谓集中式协作指的是所有无人机子系统与中心处理器 (如云平台) 通信, 将所采集数据上传至中心节点, 并接收中心处理器新的控制指令. 此类算法需建立远距离通信网络连接所有无人机, 对中心节点的算力要求较高, 因此扩展能力 (scalability) 较差. 分布式算法指的是系统中不含中心处理器, 所有子系统进行独立自主的信息处理与决策, 但子系统与子系统间可通过局部信道进行通信 (例如, 短距离蓝牙通信). 因此, 分布式系统扩展性较强, 适用于大规模污染事故<sup>[110]</sup>.

在集中式协作方面, 文献 [105] 利用平台协作进行浓度数据采集, 并通过 CUSUM 滤波器估计边界范围为 Bang-bang 控制提供了较为准确的参考信号. 大量实验表明通过多平台协作得到的边界估计和跟踪效果远优于单平台结果. Wu 和 Zhang<sup>[108]</sup> 利用卡尔曼滤波对所有无人平台数据采集进行融合, 估计编队中心梯度和浓度信息, 并利用梯度下降算法设计了编队控制策略. Sun<sup>[112]</sup> 在其博士论文中将经典的鲁棒和自适应算法引入到编队控制, 并实现了有效的边界跟踪. 文献 [106, 107] 将多平台分为两组, 分别用于边界估计和跟踪. 估计算法采用了集群优化算法 (swarm optimisation), 跟踪采用了比例控制器. 该组合方法的潜在缺点是计算复杂度高, 集群优化算法的实时性有待考量. 此外, 无人平台分组执行不同任务易造成资源浪费, 且增加了通信负荷. 从前文提及的 DCEE 的角度来说, 边界估计属于探索环境的范畴, 而边界跟踪是控制利用的范畴, 实现两者平衡在边界估计和跟踪方面也有重要意义.

Li 等<sup>[109]</sup> 采用分布式编队控制算法实现对动态污染边界的跟踪. 为估计动态扩散边界, 文中提出基于协作观测器的反馈算法. 该算法采用了距离最近子系统间的通信, 每个子系统有独立的控制决策和估计单元, 避免了对中心处理器的要求. Euler 等<sup>[110]</sup> 提出了以分布式 MPC 控制为基础的多无人机协同边界跟踪. 该算法首先将搜寻空间离散网格化, 考虑到邻近单元浓度分布的相似性, 故提出了“影响域” (impact area) 的概念构建邻近单元的浓度分布. 为增强采样效率, 文中将信息增益的概念引入到自适应采样算法中, 该增益旨在从候选采样点中选取能够最大化方差的位置. 需特别指出的是, 尽管文献 [110] 采用了分布式多无人机协同跟踪, 但主动学习的概念 (信息路径规划) 结合大规模离散化地图和非线性 MPC, 将大大增加计算复杂度.

## 5 展望

从目前无人机搜寻算法的发展趋势来看,信息与控制的融合正逐步取代传统的单极发展模式<sup>[113]</sup>.信息驱动类算法强调对未知环境的感知、预测与推理,而控制算法驱动无人机完成相应的搜寻任务.两者的有机结合与平衡已经在仿真和实验中取得了瞩目的结果.下面重点介绍对未来研究方向的想法.对于污染源跟踪,双重控制提供了一个清晰视角来剖析相关文献的本质.根据双重控制的框架,未来算法研究可聚焦于以下三方面.

- 多步滚动双重控制 (multi-stage receding horizon dual control). 文献 [24] 已给出多步双重控制的算法框架,但由于贝叶斯估计与路径规划深度耦合,采用多步预测将大大增加计算负荷.目前,无论是控制还是信息驱动算法大多采用单步预测,或可称为短视算法 (myopic algorithms),这限制了两类算法发现多步最优的能力.文献 [8] 提出了多步滚动 Infotaxis 算法,该算法结合了 RRT (rapidly-exploring random trees) 路径生成技术,多组对比实验都优于 Infotaxis. 实现多步双重控制的关键在于选用或设计更高效的估计算法.传统参数化滤波器 (如 EKF, UKF) 虽计算高效,但过度依赖模型和噪声性质,而无模型滤波器 (如粒子滤波) 计算负荷过大.未来或可尝试混杂估计算法以寻求性能和计算负荷的平衡.

- 协作搜寻. 目前已有许多文献采取多无人机协作的搜寻方案<sup>[62, 63, 102, 114, 115]</sup>,但大多为信息驱动类探索搜寻.对于多平台搜寻任务,也亟需建立统一的解决方法以实现探索与利用的平衡.相较于单平台双重控制,多无人机协作还需考虑如何平衡个体与全局的关系、如何处理个体与全局的冲突以及个体间的交互与避障.为了充分实现群体智能的优势,传统集中式方法往往无法协调处理大规模网络,去中心化搜寻是未来发展的重点方向,独立子系统的计算负荷需尽可能降低.

- 对比研究. 尽管自主搜寻已出现大量仿真和实验,但相关结果一般仅关注某一个或某一类特定算法,缺少系统地对实验和分析.构建标准化的对比实验测试各类算法的性能对未来算法研究有重要的指导意义,也有助于全面理解各类算法的内在联系和区别,例如控制驱动类、信息驱动类和双重控制算法.

对于污染边界跟踪监控问题,尽管越来越多的工作开始强调边界估计的重要性,但目前对于环境的估计仍然停留在被动学习的范畴.换言之,无人机路径决策并未将估计性能引入决策指标,其决策仅体现了边界跟踪的效果 (如 Bang-bang 控制、滑模控制等).基于前述观察和双重控制在污染源定位的进展,未来研究或可聚焦于以下几方面.

- 信息驱动类边界探寻. 如表 2 所示,许多研究通过环境学习构建边界信息,以此为无人机跟踪提供参考,估计的准确性间接决定了跟踪的性能.主动学习的方法 (信息驱动) 可引入到边界估计, Rhodes 等<sup>[116]</sup> 将信息路径规划算法用于边界探索和污染分布地图重建.后续工作可考虑将其他信息增益指标和学习算法结合,完善基于主动探索的边界跟踪.对于动态边界跟踪问题,充分利用主动学习的自适应性有助于提升估计的准确性,并强化边界跟踪性能.

- 双重控制. 将控制目标与估计目标统一、并建立两者平衡对实现自主边界跟踪有重要意义,目前这方面的研究较为欠缺,可参考文献 [24] 对该问题进行重新描述.边界跟踪的控制和信息目标与污染源定位有较大区别,相应的算法设计需根据问题属性进行调整、改进.鉴于覆盖类控制的特定需求,双重控制在多无人机协作边界跟踪的理论和实践方面都值得关注.

- 分布式边界跟踪. 由于边界跟踪需要覆盖特定的污染区域,在实际问题中污染范围往往较为广阔 (如海洋油田泄漏),传统集中式方法难以处理多无人机平台间的通信、调度、协作.目前,在多智能体系统的分布式控制、优化、决策方面已涌现出大量研究结果<sup>[117]</sup>,可用于解决大范围群体网络化边

界跟踪. 同时, 未来研究可将双重控制的思想引入到分布式边界跟踪算法, 并全面对比各类算法在分布式框架下的性能.

## 6 结论

本文将污染物监控大致分为两类: 污染源搜寻和污染边界跟踪, 着重回顾了实现无人机自主搜寻的相关算法. 无论是边界跟踪还是污染源定位, 相关算法的发展趋势都更加注重对环境的建模和学习, 污染源定位和边界跟踪的性能都很大程度上依赖于环境及污染分布估计的表现. 这一趋势本质上受到计算能力提升和相关学习算法发展的大力推动. 在污染源定位方面, 相关文献或采用控制驱动或采用信息驱动两类搜寻策略, 前者在于利用现有信息驱使无人机靠近估计污染源位置, 后者在于通过探索环境获得更好的源参数信息. 从双重控制的角度来说, 前者属于纯利用策略, 后者则为纯探索策略<sup>[24]</sup>. 根据该框架, 最新的研究聚焦在实现两者的平衡, 这对于时效性要求较高的搜寻任务至关重要, 相关理论和实验结果也揭示出双重控制在污染源参数估计和跟踪方面的巨大潜力<sup>[44]</sup>. 不同于污染源定位, 边界跟踪问题更加强调对区域的覆盖, 单个无人机平台往往无法实现有效监控, 因此许多研究采用多平台协调跟踪. 本文分别对单个和群组边界跟踪算法进行了回顾, 我们注意到目前的研究趋势与污染源定位类似, 更加注重对环境和污染浓度分布的估计和重构. 这与现代化智能控制的发展趋势相契合, 智能系统更加强调信息与控制的深度耦合, 更加注重学习、推理和决策的智能化.

## 参考文献

- 1 Hutchinson M, Oh H, Chen W H. A review of source term estimation methods for atmospheric dispersion events using static or mobile sensors. *Inf Fusion*, 2017, 36: 130–148
- 2 Hutchinson M, Liu C, Thomas P, et al. Unmanned aerial vehicle-based hazardous materials response: information-theoretic hazardous source search and reconstruction. *IEEE Robot Automat Mag*, 2019, 27: 108–119
- 3 Pajares G. Overview and current status of remote sensing applications based on unmanned aerial vehicles (UAVs). *Photogram Engng Rem Sens*, 2015, 81: 281–330
- 4 Villa T F, Gonzalez F, Miljevic B, et al. An overview of small unmanned aerial vehicles for air quality measurements: present applications and future prospectives. *Sensors*, 2016, 16: 1072
- 5 Burgués J, Marco S. Environmental chemical sensing using small drones: a review. *Sci Total Environ*, 2020, 748: 141172
- 6 Azzollini I A, Mimmo N, Gentilini L, et al. UAV-based search and rescue in avalanches using ARVA: an extremum seeking approach. 2021. ArXiv:2106.14514
- 7 Rhodes C, Liu C, Westoby P, et al. Autonomous search of an airborne release in urban environments using informed tree planning. 2021. ArXiv:2109.03542
- 8 An S, Park M, Oh H. Receding-horizon RRT-infotaxis for autonomous source search in urban environments. *Aerosp Sci Technol*, 2022, 120: 1270–9638
- 9 Bourgault F, Furukawa T, Durrant-Whyte H F. Optimal search for a lost target in a Bayesian world. In: *Field Serv Robot*. Berlin: Springer, 2003. 209–222
- 10 Pieri D, Diaz J A, Bland G, et al. In situ observations and sampling of volcanic emissions with NASA and UCR unmanned aircraft, including a case study at Turrialba Volcano, Costa Rica. *Geol Soc London Spec Publ*, 2013, 380: 321–352
- 11 Astuti G, Giudice G, Longo D, et al. An overview of the “Volcan Project”: an UAS for exploration of volcanic environments. *J Intell Robot Syst*, 2009, 54: 471–494
- 12 Salamí E, Barrado C, Pastor E. UAV flight experiments applied to the remote sensing of vegetated areas. *Remote Sens*, 2014, 6: 11051–11081

- 13 Varotto L, Cenedese A, Cavallaro A. Active sensing for search and tracking: a review. 2021. ArXiv:2112.02381
- 14 Gelenbe E, Schmajuk N, Staddon J, et al. Autonomous search by robots and animals: a survey. *Robot Autonom Syst*, 1997, 22: 23–34
- 15 Zhang X-J, Zhang M-L, Meng Q-H, et al. A gas/odor source localization strategy for mobile robot based on animal predatory behavior. *Robot*, 2008, 30: 265–272 [张小俊, 张明路, 孟庆浩, 等. 一种基于动物捕食行为的机器人气味源定位策略. *机器人*, 2008, 30: 268–272]
- 16 Bayat B, Crasta N, Crespi A, et al. Environmental monitoring using autonomous vehicles: a survey of recent searching techniques. *Curr Opin Biotechnol*, 2017, 45: 76–84
- 17 Chen X-X, Huang J. Odor source localization algorithms on mobile robots: a review and future outlook. *Robot Autonom Syst*, 2019, 112: 123–136
- 18 Shu L, Mukherjee M, Xu X, et al. A survey on gas leakage source detection and boundary tracking with wireless sensor networks. *IEEE Access*, 2016, 4: 1700–1715
- 19 Floreano D, Wood R J. Science, technology and the future of small autonomous drones. *Nature*, 2015, 521: 460–466
- 20 Luo B. Research on the key problems of odor source searching using a flying robot in time-variant airflow environments. Dissertation for Ph.D. Degree. Tianjin: Tianjin University, 2017 [罗冰. 时变气流环境下飞行机器人搜寻气味源关键问题研究. 博士学位论文. 天津: 天津大学, 2017]
- 21 Holmes N S, Morawska L. A review of dispersion modelling and its application to the dispersion of particles: an overview of different dispersion models available. *Atmos Environ*, 2006, 40: 5902–5928
- 22 Vergassola M, Villermaux E, Shraiman B I. ‘Infotaxis’ as a strategy for searching without gradients. *Nature*, 2007, 445: 406–409
- 23 Hutchinson M, Oh H, Chen W H. Entrotaxis as a strategy for autonomous search and source reconstruction in turbulent conditions. *Inf Fusion*, 2018, 42: 179–189
- 24 Chen W-H, Rhodes C, Liu C. Dual control for exploitation and exploration (DCEE) in autonomous search. *Automatica*, 2021, 133: 109851
- 25 Ristic B, Skvortsov A, Gunatilaka A. A study of cognitive strategies for an autonomous search. *Inf Fusion*, 2016, 28: 1–9
- 26 Mascarich F, Wilson T, Papachristos C, et al. Radiation source localization in GPS-denied environments using aerial robots. In: *Proceedings of IEEE International Conference on Robotics and Automation (ICRA)*, 2018. 6537–6544
- 27 Chen X, Huang J. Combining particle filter algorithm with bio-inspired anemotaxis behavior: a smoke plume tracking method and its robotic experiment validation. *Measurement*, 2020, 154: 107482
- 28 Li Z, Chen W-H, Yang J. Concurrent learning based dual control for exploration and exploitation in autonomous search. 2021. ArXiv:2108.08062
- 29 Jiang X, Li S, Luo B, et al. Source exploration for an under-actuated system: a control-theoretic paradigm. *IEEE Trans Contr Syst Technol*, 2019, 28: 1100–1107
- 30 Reinhart R, Dang T, Hand E, et al. Learning-based path planning for autonomous exploration of subterranean environments. In: *Proceedings of IEEE International Conference on Robotics and Automation (ICRA)*, 2020. 1215–1221
- 31 Patten T, Martens W, Fitch R. Monte Carlo planning for active object classification. *Auton Robot*, 2018, 42: 391–421
- 32 Denzler J, Zobel M, Niemann H. Information theoretic focal length selection for real-time active 3-D object tracking. In: *Proceedings 9th IEEE International Conference on Computer Vision*, 2003
- 33 Meera A A, Popović M, Millane A, et al. Obstacle-aware adaptive informative path planning for UAV-based target search. In: *Proceedings of International Conference on Robotics and Automation (ICRA)*, 2019. 718–724
- 34 Jiang P, Meng Q-H, Zeng M, et al. A novel visual search method for gas leakage source based on mobile robot. *Robot*, 2009, 31: 397–403 [蒋萍, 孟庆浩, 曾明, 等. 一种新的移动机器人气体泄漏源视觉搜寻方法. *机器人*, 2009, 31: 397–403]
- 35 Hutchinson M, Liu C, Chen W-H. Source term estimation of a hazardous airborne release using an unmanned aerial vehicle. *J Field Robot*, 2019, 36: 797–817



- 36 Hutchinson M, Liu C, Chen W H. Information-based search for an atmospheric release using a mobile robot: algorithm and experiments. *IEEE Trans Contr Syst Technol*, 2018, 27: 2388–2402
- 37 Stock J B, Baker M. Chemotaxis. In: *Encyclopedia of Microbiology*. Pittsburgh: Academic Press, 2009. 71–79
- 38 Liu S-J, Krstic M. Stochastic source seeking for nonholonomic unicycle. *Automatica*, 2010, 46: 1443–1453
- 39 Ramirez-Llanos E, Martinez S. Stochastic source seeking for mobile robots in obstacle environments via the SPSA method. *IEEE Trans Automat Contr*, 2018, 64: 1732–1739
- 40 Guay M, Zhang T. Adaptive extremum seeking control of nonlinear dynamic systems with parametric uncertainties. *Automatica*, 2003, 39: 1283–1293
- 41 Barbieri C, Cocco S, Monasson R. On the trajectories and performance of infotaxis, an information-based greedy search algorithm. *Europhys Lett*, 2011, 94: 20005
- 42 Khodayi-mehr R, Aquino W, Zavlanos M M. Model-based active source identification in complex environments. *IEEE Trans Robot*, 2019, 35: 633–652
- 43 Hollinger G A, Englot B, Hover F S, et al. Active planning for underwater inspection and the benefit of adaptivity. *Int J Robot Res*, 2013, 32: 3–18
- 44 Rhodes C, Liu C, Chen W H. Autonomous source term estimation in unknown environments: from a dual control concept to UAV deployment. *IEEE Robot Autom Lett*, 2022, 7: 2274–2281
- 45 Ghassemi P, Chowdhury S. An extended bayesian optimization approach to decentralized swarm robotic search. *J Comput Inf Sci Eng*, 2020, 20: 051003
- 46 Marchant R, Ramos F. Bayesian optimisation for intelligent environmental monitoring. In: *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2012. 2242–2249
- 47 Albani R A S, Albani V V L, Migon H S, et al. Uncertainty quantification and atmospheric source estimation with a discrepancy-based and a state-dependent adaptative MCMC. *Environ Pollution*, 2021, 290: 118039
- 48 Kumar P R. Convergence of adaptive control schemes using least-squares parameter estimates. *IEEE Trans Automat Contr*, 1990, 35: 416–424
- 49 Guo L. On critical stability of discrete-time adaptive nonlinear control. *IEEE Trans Automat Contr*, 1997, 42: 1488–1499
- 50 Kreucher C, Hero A O, Kastella K. A comparison of task driven and information driven sensor management for target tracking. In: *Proceedings of the 44th IEEE Conference on Decision and Control*, 2005. 4004–4009
- 51 Meng Q-H, Li F, Zhang M-L, et al. Study on realization method of multi-robot active olfaction in turbulent plume environments. *Acta Autom Sin*, 2008, 34: 1281–1290 [孟庆浩, 李飞, 张明路, 等. 湍流烟羽环境下多机器人主动嗅觉实现方法研究. *自动化学报*, 2008, 34: 1281–1290]
- 52 Howard T, Pivtoraiko M, Knepper R A, et al. Model-predictive motion planning: several key developments for autonomous mobile robots. *IEEE Robot Automat Mag*, 2014, 21: 64–73
- 53 Nguyen H, Kamel M, Alexis K, et al. Model predictive control for micro aerial vehicles: a survey. 2020. ArXiv:2011.11104
- 54 Ortega R, Nikiforov V, Gerasimov D. On modified parameter estimators for identification and adaptive control. A unified framework and some new schemes. *Annu Rev Control*, 2020, 50: 278–293
- 55 Mahony R, Kumar V, Corke P. Multirotor aerial vehicles: modeling, estimation, and control of quadrotor. *IEEE Robot Automat Mag*, 2012, 19: 20–32
- 56 Chen W H, Ballance D J, O'Reilly J. Model predictive control of nonlinear systems: computational burden and stability. *IEE Proc-Control Theor Appl*, 2000, 147: 387–394
- 57 Ariyur K B, Krstic M. *Real-Time Optimization by Extremum-Seeking Control*. Hoboken: John Wiley & Sons, 2003
- 58 Krstić M, Wang H H. Stability of extremum seeking feedback for general nonlinear dynamic systems. *Automatica*, 2000, 36: 595–601
- 59 Tan Y, Moase W H, Manzie C, et al. Extremum seeking from 1922 to 2010. In: *Proceedings of the 29th Chinese Control Conference*, 2010. 14–26
- 60 Cochran J, Siranosian A, Ghods N, et al. 3-D source seeking for underactuated vehicles without position measurement.

- IEEE Trans Robot, 2009, 25: 117–129
- 61 Bıyık E, Arcak M. Gradient climbing in formation via extremum seeking and passivity-based coordination rules. *Asian J Control*, 2008, 10: 201–211
- 62 Soares J M, Aguiar A P, Pascoal A M, et al. A distributed formation-based odor source localization algorithm-design, implementation, and wind tunnel evaluation. In: *Proceedings of IEEE International Conference on Robotics and Automation (ICRA)*, 2015. 1830–1836
- 63 Park M, Oh H. Cooperative information-driven source search and estimation for multiple agents. *Inf Fusion*, 2020, 54: 72–84
- 64 Poveda J I, Benosman M, Teel A R, et al. Robust coordinated hybrid source seeking with obstacle avoidance in multivehicle autonomous systems. *IEEE Trans Automat Contr*, 2021, 67: 706–721
- 65 Li Z, You K, Song S. Cooperative source seeking via networked multi-vehicle systems. *Automatica*, 2020, 115: 108853
- 66 Qu X-M, Liu T, Tan W-R. Multi-source passive localization via multiple unmanned aerial vehicles. *Sci Sin Inform*, 2019, 49: 570–584 [屈小媚, 刘韬, 谈文蓉. 基于多无人机协作的多目标无源定位算法. *中国科学: 信息科学*, 2019, 49: 570–584]
- 67 Cover T M. *Elements of Information Theory*. Hoboken: John Wiley & Sons, 1999
- 68 Miller L M, Silverman Y, MacIver M A, et al. Ergodic exploration of distributed information. *IEEE Trans Robot*, 2015, 32: 36–52
- 69 Wang H. Minimum entropy control of non-Gaussian dynamic stochastic systems. *IEEE Trans Automat Contr*, 2002, 47: 398–403
- 70 Zhao Y, Chen B, Zhu Z, et al. Entrotaxis-jump as a hybrid search algorithm for seeking an unknown emission source in a large-scale area with road network constraint. *Expert Syst Appl*, 2020, 157: 113484
- 71 Moraud E M, Martinez D. Effectiveness and robustness of robot infotaxis for searching in dilute conditions. *Front Neurobot*, 2010, 4: 1
- 72 Alpcan T, Shames I. An information-based learning approach to dual control. *IEEE Trans Neural Netw Learn Syst*, 2015, 26: 2736–2748
- 73 Klenske E D, Hennig P. Dual control for approximate bayesian reinforcement learning. *J Mach Learn Res*, 2016, 17: 4354–4383
- 74 Guha A, Annaswamy A. MRAC-RL: a framework for on-line policy adaptation under parametric model uncertainty. 2020. ArXiv:2011.10562
- 75 Jeong H, Hassani H, Morari M, et al. Learning to track dynamic targets in partially known environments. 2020. ArXiv:2006.10190
- 76 Fisac J F, Akametalu A K, Zeilinger M N, et al. A general safety framework for learning-based control in uncertain robotic systems. *IEEE Trans Automat Contr*, 2018, 64: 2737–2752
- 77 Li Y. Deep reinforcement learning: an overview. 2017. ArXiv:1701.07274
- 78 Settles B. *Active Learning Literature Survey*. Computer Sciences Technical Report 1648, University of Wisconsin–Madison, 2009
- 79 Silver D, Schrittwieser J, Simonyan K, et al. Mastering the game of Go without human knowledge. *Nature*, 2017, 550: 354–359
- 80 Chen W-H. Perspective view of autonomous control in unknown environment: dual control for exploitation and exploration vs reinforcement learning. *Neurocomputing*, 2022, 497: 50–63
- 81 Feldbaum A A. Dual control theory I. *Avtomatika i Telemekhanika*, 1960, 21: 1240–1249
- 82 Feldbaum A A. Dual control theory II. *Avtomatika i Telemekhanika*, 1960, 21: 1453–1464
- 83 Feldbaum A A. Dual control theory III. *Avtomatika i Telemekhanika*, 1961, 22: 3–16
- 84 Feldbaum A A. Dual control theory IV. *Avtomatika i Telemekhanika*, 1961, 22: 129–142
- 85 Bar-Shalom Y, Tse E. Dual effect, certainty equivalence, and separation in stochastic control. *IEEE Trans Automat Contr*, 1974, 19: 494–500
- 86 Bertsekas D. *Dynamic Programming and Optimal Control*, vol. 1. Nashua: Athena Scientific, 2012

- 87 Bertsekas D P. *Nonlinear Programming*. 3rd ed. Nashua: Athena Scientific, 2016
- 88 Wittenmark B. Adaptive dual control methods: an overview. *IFAC Proc Volumes*, 1995, 28: 67–72
- 89 Heirung T A N, Ydstie B E, Foss B. Dual adaptive model predictive control. *Automatica*, 2017, 80: 340–348
- 90 Mesbah A. Stochastic model predictive control with active uncertainty learning: a survey on dual control. *Annu Rev Control*, 2018, 45: 107–117
- 91 Hoy M, Matveev A S, Savkin A V. Algorithms for collision-free navigation of mobile robots in complex cluttered environments: a survey. *Robotica*, 2015, 33: 463–497
- 92 Towler J, Krawiec B, Kochersberger K. Radiation mapping in post-disaster environments using an autonomous helicopter. *Remote Sens*, 2012, 4: 1995–2015
- 93 Brink J. Boundary tracking and estimation of pollutant plumes with a mobile sensor in a low-density static sensor network. *Urban Clim*, 2015, 14: 383–395
- 94 Mellucci C, Menon P P, Edwards C, et al. Experimental validation of boundary tracking using the suboptimal sliding mode algorithm. In: *Proceedings of 2017 American Control Conference (ACC)*, 2017. 4878–4883
- 95 Menon P P, Edwards C, Shtessel Y B, et al. Boundary tracking using a suboptimal sliding mode algorithm. In: *Proceedings of the 53rd IEEE Conference on Decision and Control*, 2014. 5518–5523
- 96 Matveev A S, Teimoori H, Savkin A V. Method for tracking of environmental level sets by a unicycle-like vehicle. *Automatica*, 2012, 48: 2252–2261
- 97 Dong F, You K. The isoline tracking in unknown scalar fields with concentration feedback. *Automatica*, 2021, 133: 109779
- 98 Zhang C, Pei H. Oil spills boundary tracking using universal Kriging and model predictive control by UAV. In: *Proceeding of the 11th World Congress on Intelligent Control and Automation*, 2014. 633–638
- 99 Jiang X, Li S. Plume front tracking in unknown environments by estimation and control. *IEEE Trans Ind Inf*, 2018, 15: 911–921
- 100 Susca S, Bullo F, Martinez S. Monitoring environmental boundaries with a robotic sensor network. *IEEE Trans Contr Syst Technol*, 2008, 16: 288–296
- 101 Sun T, Pei H, Pan Y, et al. Robust adaptive neural network control for environmental boundary tracking by mobile robots. *Int J Robust Nonlinear Control*, 2013, 23: 123–136
- 102 Kemp M, Bertozzi A L, Marthaler D. Multi-UUV perimeter surveillance. In: *Proceedings of IEEE/OES Autonomous Underwater Vehicles*, 2004. 102–107
- 103 Jin Z, Bertozzi A L. Environmental boundary tracking and estimation using multiple autonomous vehicles. In: *Proceedings of the 46th IEEE Conference on Decision and Control*, 2007. 4918–4923
- 104 Hsieh C H, Jin Z, Marthaler D, et al. Experimental validation of an algorithm for cooperative boundary tracking. In: *Proceedings of the American Control Conference*, 2005. 1078–1083
- 105 Joshi A, Ashley T, Huang Y R, et al. Experimental validation of cooperative environmental boundary tracking with on-board sensors. In: *Proceedings of American Control Conference*, 2009. 2630–2635
- 106 Menon P P, Ghose D. Boundary mapping of 3-dimensional regions. In: *Proceedings of American Control Conference*, 2013. 2984–2989
- 107 Menon P P, Ghose D. Simultaneous source localization and boundary mapping for contaminants. In: *Proceedings of American Control Conference (ACC)*, 2012. 4174–4179
- 108 Wu W, Zhang F. Cooperative exploration of level surfaces of three dimensional scalar fields. *Automatica*, 2011, 47: 2044–2051
- 109 Li S, Guo Y, Bingham B. Multi-robot cooperative control for monitoring and tracking dynamic plumes. In: *Proceedings of IEEE International Conference on Robotics and Automation (ICRA)*, 2014. 67–73
- 110 Euler J, Horn A, Haumann D, et al. Cooperative n-boundary tracking in large scale environments. In: *Proceedings of IEEE 9th International Conference on Mobile Ad-Hoc and Sensor Systems (MASS 2012)*, 2012. 1–6
- 111 Mesbah A. Stochastic model predictive control: An overview and perspectives for future research. *IEEE Control Syst Mag*, 2016, 36: 30–44

- 112 Sun T R. Environmental boundary tracking control for mobile robots. Dissertation for Ph.D. Degree. Guangzhou: South China University of Technology, 2011 [孙太任. 机器人对环境边界的跟踪控制. 博士学位论文. 广州: 华南理工大学, 2011]
- 113 Zheng Z M, Lv J H, Wei W, et al. Refined intelligence theory: artificial intelligence regarding complex dynamic objects. *Sci Sin Inform*, 2021, 51: 678–690 [郑志明, 吕金虎, 韦卫, 等. 精准智能理论: 面向复杂动态对象的人工智能. *中国科学: 信息科学*, 2021, 51: 678–690]
- 114 Koohifar F, Guvenc I, Sichert M L. Autonomous tracking of intermittent RF source using a UAV swarm. *IEEE Access*, 2018, 6: 15884–15897
- 115 Ristic B, Gilliam C, Moran W, et al. Decentralised multi-platform search for a hazardous source in a turbulent flow. *Inf Fusion*, 2020, 58: 13–23
- 116 Rhodes C, Liu C, Chen W-H. Informative path planning for gas distribution mapping in cluttered environments. In: *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2020. 6726–6732
- 117 Li Z, Dong Z, Liang Z, et al. Surrogate-based distributed optimisation for expensive black-box functions. *Automatica*, 2021, 125: 109407

## Review of UAV-based autonomous search algorithms for hazardous sources

Zhongguo LI, Wen-Hua CHEN\* & Cunjia LIU

*Department of Aeronautical and Automotive Engineering, Loughborough University, Loughborough LE11 3TU, UK*

\* Corresponding author. E-mail: w.chen@lboro.ac.uk

**Abstract** Autonomous search plays an important role in a wide range of applications, such as environmental protection, industrial monitoring, and search and rescue. Sensor platforms based on unmanned aerial vehicles (UAVs) have been regarded as one of the most efficient methods for achieving fully autonomous searching because of their superior flexibility and motility. This paper reviews existing algorithms for two important subjects, i.e., source seeking and boundary tracking, based on which their design principles are revealed. Although control-oriented and information-driven approaches are the two mainstream state-of-the-art methods for source seeking, their hybrids have been poorly studied. Regarding boundary tracking, existing solutions can be allocated into two classes: single-agent tracking and multiagent cooperative tracking. From the perspective of dual control for exploration and exploitation, this review outlines the fundamental insights into algorithm design and presents several visions for future research.

**Keywords** autonomous search, UAV, source seeking, boundary tracking, dual control, exploration and exploitation