IEEE Transactions on Aerospace and Electronic Systems, Volume 59, Issue 1, February 2023, pp. 591-609
DOI:10.1109/TAES.2022.3184923

Receding Horizon-Based Infotaxis with Random Sampling for Source Search and Estimation in Complex Environments

Minkyu Park[a], Pawel Ladosz[a], Jongyun Kim[b], and Hyondong Oh[a,*], Senior Member, IEEE

***Abstract*—This paper proposes a receding horizon-based information-theoretic source search and estimation strategy for a mobile sensor in an urban environment in which an invisible harmful substance is released into the atmosphere. The mobile sensor estimates the source term including its location and release rate by using sensor observations based on Bayesian inference. The sampling-based sequential Monte Carlo method, particle filter, is employed to estimate the source term state in a highly nonlinear and stochastic system. Infotaxis, the information-theoretic gradient-free search strategy is modified to find the optimal search path that maximizes the reduction of the entropy of the source term distribution. In particular, receding horizon Infotaxis is introduced to avoid falling into the local optima and to find more successful information gathering paths in obstacle-rich urban environments. Besides, a random sampling method is introduced to reduce the computational load of the receding horizon Infotaxis for real-time computation. The random sampling method samples the predicted future measurements based on current estimation of the source term and computes the optimal search path using sampled measurements rather than considering all possible future measurements. To demonstrate the benefit of the proposed approach, comprehensive numerical simulations are performed for various conditions. The proposed algorithm increases the success rate by about 30% and reduces the mean search time by about 40% compared with the existing information-theoretic search strategy.**

***Index Terms*—Information-theoretic search, Autonomous mobile sensor management, Bayesian inference, Sequential Monte Carlo method, Dispersion modeling, Receding horizon path planning**

I. Introduction

The potential danger of accidental release of harmful chemical, biological, radiological, or nuclear (CBRN) materials into the atmosphere has significantly increased in recent years [1]. Estimating the source term, including the source location and release strength, is of primary importance for responding to emergencies. Generally, atmospheric dispersion is treated as a highly nonlinear and turbulent event with sporadic and fluctuating sensing cues. Furthermore, in many cases, CBRN materials are dispersed in the atmosphere as an invisible substance. The sparse sensing cue and the invisible nature of the substance make the source term estimation problem challenging. The traditional source term estimation method employs static sensors [2], however, static sensors have a great limitation when accidents occur in unexpected places. In particular, this method has a limited response area as the sensors should be installed before the accident. Thus, search strategies using mobile sensing agents have attracted considerable attention over the last decade, where unmanned aerial vehicles (UAVs) are frequently used as mobile sensors thanks to their flexibility and versatility [3]–[11].

Various mobile sensing strategies for source term estimation have been proposed, which can be largely categorized into: gradient climbing [12], the reactive strategy [13]–[16], and cognitive/probabilistic methods [17]–[20]. The gradient climbing method is often inspired by the behavior of bacteria searching for nutrition [12]. However it does not work well when the sensing cue is not continuous or frequently lost, e.g., in turbulent atmospheric dispersion cases. The reactive search method is inspired by behaviors of different species, such as finding nutrition, mating, and hunting. The mobile sensor moves with predefined patterns such as a zigzag or spiral path as a reaction to sensing cues [21], [22]. These methods have the advantage of being able to find the source even if the concentration does not change continuously, but they do not guarantee the optimal performance. Cognitive search strategies aim to solve the real-time source term estimation and optimal search path planning problem [8], [17]–[20], [23], [24]. The cognitive strategy uses the information theory to determine the optimal search path while estimating the source term, e.g., Infotaxis [17], [18], [20] and Entrotaxis [19], [24]; thus, they are often called an information-theoretic search. This approach is proven to be robust to sparse or intermittent sensing cues, so it is frequently used in turbulent environments on which this study focuses.

However, many cognitive search strategies decide the action greedily and many studies assume a gas dispersion situation in an open space [8], [17]–[20], [23], [24]. The accidents and disasters that require rapid response often occur in urban or industrial environments that have numerous buildings or structures in the search domain. In general, the search path generated by a one step lookahead decision making, which can be called a greedy algorithm, is inefficient in an obstacle-rich environment. Although several existing studies for source search in the obstacle-rich environment is conducted, many of them assumed known environments (i.e., positions of obstacles are given) [25], [26] and used separate obstacle avoidance algorithm in unknown environments that is not tightly integrated

[a] Department of Mechanical Engineering, Ulsan National Institute of Science and Technology, Ulsan 44919, South Korea [b] School of Aerospace, Transport and Manufacturing, Cranfield University, Cranfield, Bedfordshire, MK43 0AL, United Kingdom
(e-mail: mk.park@unist.ac.kr, pladosz@unist.ac.kr, jongyun.kim@cranfield.ac.uk, and h.oh@unist.ac.kr).
*Corresponding author.

with search strategies to obtain better information about the source [27]–[30], leading to inefficient or unsuccessful source search.

With these backgrounds in mind, this paper presents a receding horizon-based information-theoretic source search algorithm tightly combined with the obstacle avoidance feature for accurate and robust source term estimation using a mobile sensor in an unknown and complex urban environment, as illustrated in Fig. 1. The main contribution of this study is to propose the multi-step lookahead (i.e., receding horizon) search path planning for source term estimation which maximizes the entropy reduction based on Infotaxis termed as receding horizon Infotaxis (RHI). The proposed RHI strategy considers the search path by using the predicted future measurements based on the current estimation to obtain the best source information while simultaneously avoiding obstacles in urban environments.

To implement the RHI, we consider the use of two sensor models: Gaussian sensor model [31] and binary sensor model [32]. In particular, the binary sensor model provides detection or non-detection measurements. Thus, it requires the fewer number of states than that of the Gaussian sensor model that provides the continuous concentration of a gas. Note that the computational complexity increases exponentially as the number of horizon steps increases as the receding horizon method takes into account all combinations of predicted future measurement sequences at sensing points of multiple steps [33]. To address this issue, we utilize a random sampling approach that uses only a few sampled future measurements from the estimated probability distribution of the source term.

To investigate the trade off between the performance and computation time, we propose and compare different receding horizon Infotaxis (RHI) methods by incorporating the Gaussian sensor model or binary sensor model with/without the random sampling method: (i) RHI-G (receding horizon Infotaxis with the Gaussian sensor model), (ii) RHI-B (receding horizon Infotaxis with the binary sensor model), (iii) RHI-GR (receding horizon Infotaxis with the Gaussian sensor model and random sampling), and (iv) RHI-BR (receding horizon Infotaxis with the binary sensor model and random sampling). The various numerical simulations are performed to validate the effectiveness and robustness of the proposed methods in terms of search time, estimation accuracy, success rate and computation time.

The remainder of this paper is organized as follows. In Section II, the problem description, including the gas dispersion and sensor model, is presented. Section III describes the estimation method for the source term based on the sequential Monte Carlo method (i.e., particle filter). The receding horizon-based information-theoretic search path planning algorithm is presented in Section IV. Section V provides numerical simulation results and comparison studies. Lastly, conclusions and directions of future work are given in Section VI.

II. Problem Description

The mobile sensor (e.g., UAV) to search and collect sensor data is assumed to move in a horizontal grid map by selecting the best action from the feasible action set $\mathbf{u} = [\uparrow, \downarrow, \rightarrow, \leftarrow]$ on the horizontal plane at each search time step; hence, this is a sequential decision making problem. During one time step, the agent obtains the sensor data (which normally takes a few seconds for gas sensing [8]) at the current sensing position, estimates the source term, computes the best action direction (i.e., path planning), and moves to the next sensing position determined by the selected direction with a fixed distance in a constant speed. Since it is assumed that the sensing time and moving time between consecutive sensing points are fixed, the computation time for estimation and path planning per step as well as the number of time steps plays an important role for total source search time. Although more feasible actions (e.g., diagonal motions or different distances between sensing points) or trajectory planning with a variable agent speed might improve the search performance (in terms of the number of time steps), it would significantly increase the computation time. Besides, it is assumed that the mobile agent is equipped with a range sensor such as LiDAR and it can build a local obstacle map so that it could plan the long-term search path while avoiding obstacles without the prior geometrical map information. Note that we mainly consider a 2-D environment for simplicity but the developed framework could be readily applicable to a 3-D environment. In fact, we provide a representative simulation result for a 3-D environment in the numerical simulation section.

From the following, the mathematical formulation of the gas dispersion model and the sensor model will be explained: the isotropic plume model [17] is used as the dispersion model; and the Gaussian sensor model is used as the stochastic sensor model [34], [35].

A. Gas dispersion model

The isotropic plume model, which is an analytical gas dispersion model for an open space [17], is used as the dispersion model to compute the probability distribution of the source term. The plume model assumes that the gas is continuously released from the stationary point source located at $\mathbf{p}_s = [x_s,\ y_s]^\mathsf{T} \in \mathbb{R}^{2+}$ with the fixed release strength $Q_s \in \mathbb{R}^+$, and it calculates the mean gas concentration $\mu(\mathbf{p}_k;\theta)$ at the sensing location $\mathbf{p}_k = [x_k,\ y_k]^\mathsf{T}$ at k^{th} time step as:

$$\mu(\mathbf{p}_k;\theta) = \frac{Q_s}{4\pi D|\mathbf{p}_k - \mathbf{p}_s|} \exp\left[\frac{-(y_k - y_s)V}{2\zeta}\right] \cdot \exp\left[-\frac{|\mathbf{p}_k - \mathbf{p}_s|}{\lambda}\right]\Delta t, \tag{1}$$

where

$$\lambda = \sqrt{\frac{\zeta\tau}{1 + \frac{V^2\tau}{4\zeta}}}. \tag{2}$$

Here, θ represents the source term parameters including the mean wind speed V, the release strength Q_s and the source location $\mathbf{p}_s$ and other parameters to compute the concentration in Eq. (1). It is assumed that the wind blows in the $-y$ direction; thus, x represents the crosswind direction. The isotropic effective diffusion parameter ζ affects the dispersion in the crosswind direction when the gas is advected by the

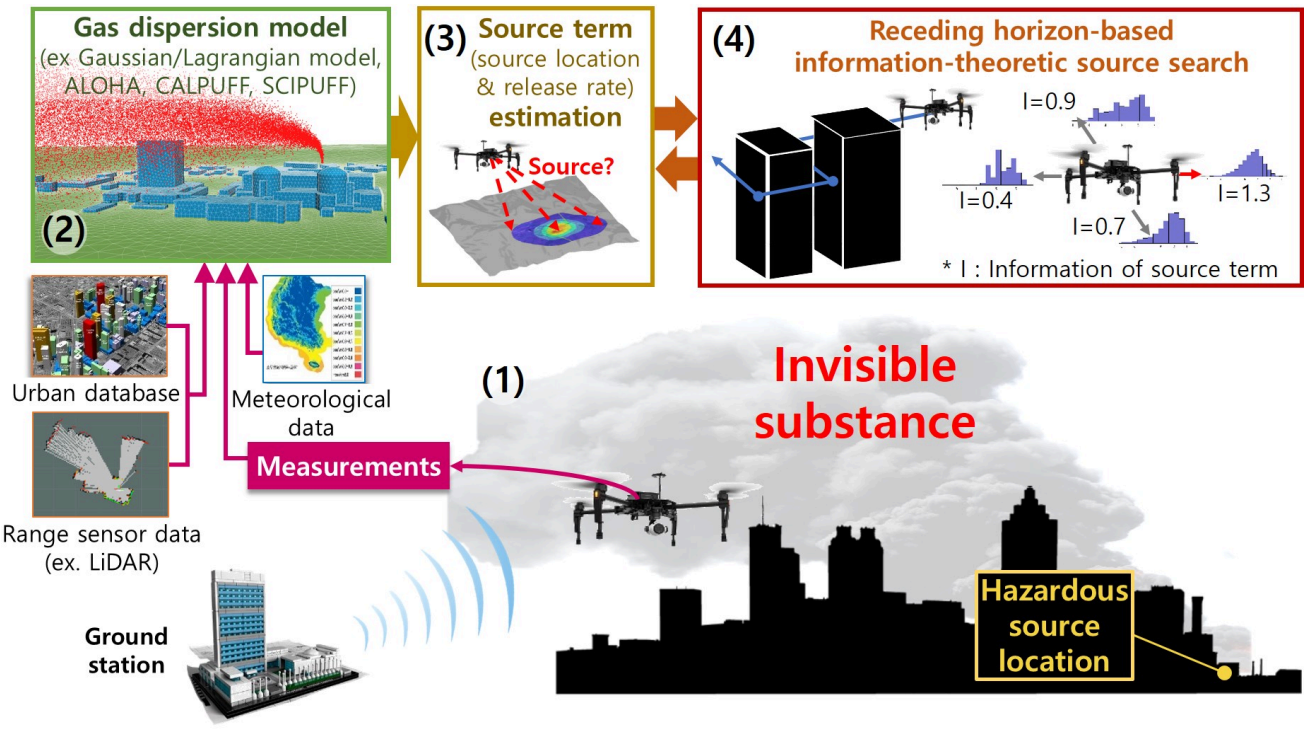


Fig. 1. Overall architecture of the source search and estimation algorithm: (1) The UAV obtains sensor measurements at sensing position; (2) External environmental data are received; (3) Estimation is conducted; and (4) The optimal search path is generated by using receding horizon-based information-theoretic source search.

mean wind velocity V. The gas particle can be maintained for the lifetime τ. Since the coordinate axis y of Eq. (1) is aligned to the direction of advection, rotation transformation is used to express gas dispersion in arbitrary directions in the inertial coordinate system for the search area as

$$\begin{bmatrix} x_i \\ y_i \\ z_i \end{bmatrix} = \begin{bmatrix} \cos(\phi) & -\sin(\phi) & 0 \\ \sin(\phi) & \cos(\phi) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \tag{3}$$

where the $[x_i, y_i, z_i]^\mathsf{T}$ and $[x, y, z]^\mathsf{T}$ are the relative distance from the source in inertial and dispersion frames, respectively, and ϕ represents the wind direction in the inertial coordinate.

B. Sensor models

The mobile sensor is equipped with a gas sensor that measures the gas concentration. In this study, we compare the performance of the source term estimation using two different gas sensor models: the Gaussian sensor model and the binary sensor model. The Gaussian sensor model directly provides the concentration of the surrounding gas which contains the stochastic noise while the binary sensor model only distinguishes detection and non-detection cases. Note that, two main parts of the search algorithm utilize the sensor model: the particle filter and receding horizon-based search path planning. The binary sensor model is adopted to reduce the computational burden only for the path planning part, which will be explained in detail later.

1) Gaussian sensor model: One of the most conservative choices for the sensor model is Gaussian-based models [31], [34]–[36]. Since the actual concentration detected by the sensor follows the stochastic process, the Gaussian sensor model can be employed as the sensor model to estimate source term and predict the future measurement. The sensor measurement $c_k \in \mathbb{R}^+$ is stochastically determined by the background concentration $\mu'_k \in \mathbb{R}^+$. The background gas concentration is easily influenced by the environment including the local atmospheric temperature and wind disturbance among others. Those environmental factors cannot be easily determined exactly due to frequently changing or fluctuating characteristics. Even though the analytical plume dispersion models calculate the time-averaged concentration of the gas $\mu_k \in \mathbb{R}^+$, only the current temporal concentration μ'_k affects the sensor measurement; thus, we generate the temporal concentration μ'_k from the analytic concentration μ_k from Eq. (1) by following the Gaussian distribution. Then the sensor measurement c_k is represented as:

$$c_k = \mu'_k + v_{sen} = (\mu_k + v_{env}) + v_{sen}, \tag{4}$$

where,

$$\begin{aligned} v_{env} &\sim \mathcal{N}\left(0, \sigma^2_{env}\right), \\ v_{sen} &\sim \mathcal{N}\left(0, \sigma^2_{sen}\right). \end{aligned} \tag{5}$$

It is assumed that the standard deviation of the environmental noise, σ_{env}, is a constant, and the sensing noise is proportional to the current background concentration, that is, $\sigma_{sen} \propto \mu'_k$. Thus, the probabilities of obtaining a certain measurement

given a source term are calculated as follows.

$$p(\mu'_k|\theta) = \frac{1}{\sigma_{env}\sqrt{2\pi}} \exp\left[-\frac{v_{env}^2}{2\sigma_{env}^2}\right], \\ p(c_k|\mu'_k,\theta) = \frac{1}{\sigma_{sen}\sqrt{2\pi}} \exp\left[-\frac{v_{sen}^2}{2\sigma_{sen}^2}\right], \tag{6}$$

where the source term vector θ is used to calculate the mean concentration μ obtained by Eq. (1).

The sample maps of the time-averaged and the temporal concentration dispersion are presented in Fig. 2(a-b), and the sample map of the sensor measurement according to the current temporal concentration is shown in Fig. 2(c). The sample maps are generated by following parameters: $Q_s = 2000mg/s$, $\zeta = 10m^2/s$, $\tau = 1000s$, $V = 2m/s$, $\Delta t = 1s$ $\phi = 220°$, $\sigma_{env} = 0.4mg/m^3$, and $\sigma_{sen} = 0.2 \cdot \mu'_k(\mathbf{p}_k)mg/m^3$. Note that the sensor parameter is a parameter that should be tuned, the chosen number in this paper is empirically tuned based on the observations of the custom MOX sensor.

As shown in Fig. 2(b), the gas dispersion is noisy, sparse and frequently absent in the stochastic gas dispersion environment. The actual sensor measurement presented in Fig. 2(c) is more noisy than Fig. 2(b). Particularly, the mobile sensor has a significantly high noise near the source since the modeled sensor noise is affected by concentration at each location. It can be observed that the contour near the source is much more unstable in Fig. 2(c) compared with Fig. 2(a-b). This kind of noise is one of the most significant factors, making it difficult to use the conventional gradient climbing method for source search in turbulent environments [12], [17].

2) Binary sensor model: The binary sensor model uses 1 (detection) or 0 (non-detection) as a sensor measurement. Thus, the probability of a binary measurement at k^{th} time step $b_k \in [0, 1]$ is expressed as:

$$p(b_k|\theta) = \begin{cases} \beta & \text{if } b_k = 0, \\ 1-\beta & \text{if } b_k = 1, \end{cases} \tag{7}$$

which is functionally equivalent to Eq. (6). This model can be treated as a special case of the Gaussian sensor model with only two measurement values. When the current concentration c_k is higher than the threshold $\bar{b}_k$, the binary sensor considers that the gas is detected, otherwise non-detected. Thus, the probability β can be calculated by the cumulative distribution function (CDF) of the standard normal distribution $\Phi(\cdot)$ [37] as:

$$\beta = \Phi\left(\frac{\Delta b_k}{\sigma_{env}}\right), \tag{8}$$

where

$$\Delta b_k = \bar{b}_k - \mu'_k, \tag{9}$$

The threshold $\bar{b}_k$, which determines whether the sensor detects the substance or not, is adaptively changed as:

$$\bar{b}_k = \begin{cases} \lambda_b \bar{b}_{k-1} + (1-\lambda_b)c_k & \text{if } k > 1,\ c_k > \bar{b}_{k-1} \\ \bar{b}_{k-1} & \text{if } k > 1,\ c_k \leq \bar{b}_{k-1} \\ c_k & \text{if } k = 1, \end{cases} \tag{10}$$

where λ_b is a user design parameter. Note that, unlike previous studies [14], [38], the threshold is only updated when the new measurement value is greater than the current threshold, so the threshold increases monotonically. The decision making utility for the search is based on the entropy of the estimated source term (, which is explained in Section IV). Thus, this threshold update and the corresponding utility make the agent select the exploitative action (i.e., moving towards the estimated source location) when the agent can find the place containing higher concentration in the future or select the explorative action (i.e., wandering around) otherwise.

III. Source Term Estimation

The source term vector contains a set of parameters such as the source location, gas release strength, and wind speed, wind direction among many others. The list of parameters depends on dispersion models but the key parameters commonly included are the source location and release strength. To respond to a disaster involving the invisible hazardous gas, estimating the source location and release strength is of primary importance; thus, the source term vector used in this study consists of source location and release strength, $\theta = [x_s\ y_s\ Q_s]^\mathsf{T} \in \mathbb{R}^3$. We assume that the other source parameters, such as the wind velocity and direction or gas diffusivity, are known. This assumption is generally accepted in many source term estimation studies [16], [18], [19], [39], as these parameters could be determined by meteorological data and the gas properties. From the following, we introduce a method for estimating the source term vector using Bayesian inference with the particle filter approach. As the computational burden for calculating the state estimation using the current measurement is not changed by the sensor model (i.e., Gaussian or binary sensor model), we conduct the source term estimation using the Gaussian sensor model only.

A. Bayesian inference

The Bayesian framework is used to estimate the source term via sensor observations. The basic Bayesian inference estimates the source term probability density function (PDF) using the prior knowledge of the source term PDF $p(\theta)$, and the new observation c as:

$$p(\theta|c) = \frac{p(c|\theta)\, p(\theta)}{p(c)}. \tag{11}$$

The estimated PDF is sequentially updated by the Bayesian framework; thus the posterior estimation for the source term at $(k-1)^{th}$ time step, $p(\theta_{k-1}|c_{1:k-1})$, is first propagated in time with the process model to have $p(\theta_k|c_{1:k-1})$, and then it is used as new prior knowledge for k^{th} time step as:

$$p(\theta_k|c_{1:k}) = \frac{p(c_k|\theta_k)\, p(\theta_k|c_{1:k-1})}{p(c_k|c_{1:k-1})}, \tag{12}$$

where the collected measurement $c_{1:k} = [c_1(\mathbf{p}_1), \cdots, c_k(\mathbf{p}_k)]$ is a sequence of observations for each sensing position. In this study, we assume that the source has no dynamics (i.e., stationary), which means $p(\theta_k|c_{1:k-1})$ is the same as $p(\theta_{k-1}|c_{1:k-1})$. The likelihood

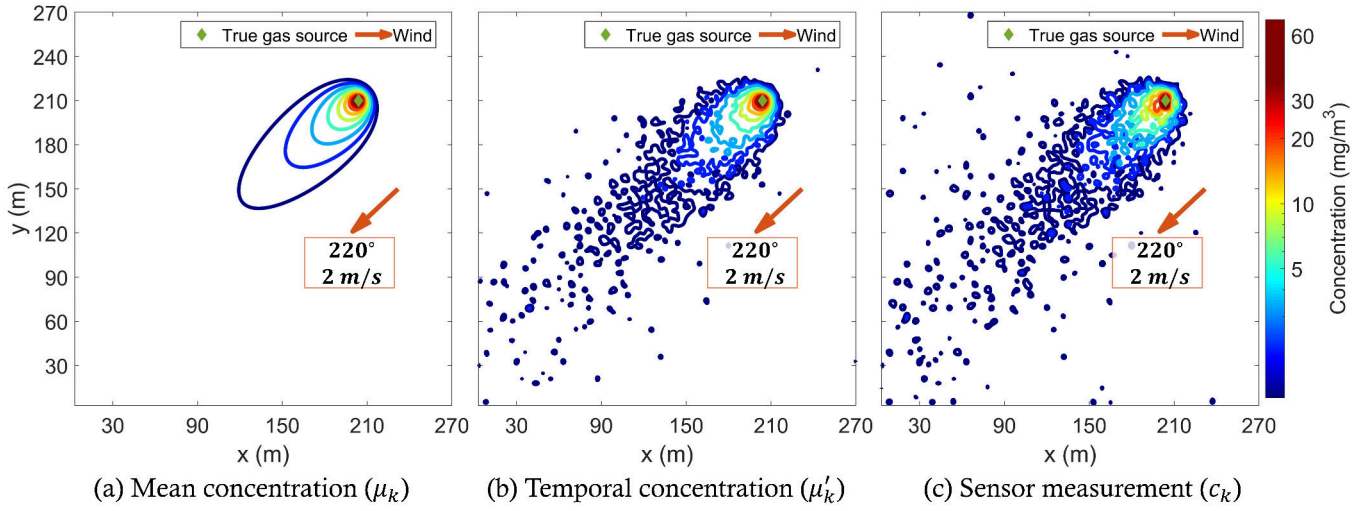


Fig. 2. The examples of gas dispersion and corresponding sensor measurement map.

$p(c_k|\theta_k)$ can be calculated by the function of the sensor model and dispersion model. The source term vector θ_k is used to calculate μ_k by using the dispersion model Eq. (1). The actual sensor measurement c_k is affected by the current background concentration μ'_k which is calculated from the time-averaged mean concentration μ_k. However, when we calculate the likelihood $p(c_k|\theta_k)$, the true current concentration μ'_k is hard to determine. To calculate the likelihood, we can use the marginal probability as:

$$p(c_k|\theta_k) = p(c_k|\mu_k) = \int_{-\infty}^{\infty} p(c_k, \mu'_k|\mu_k)\mathrm{d}\mu'_k, \tag{13}$$

where $p(c_k, \mu'_k|\mu_k)$ can be expressed as the product of $p(c_k|\mu'_k, \mu_k)$ and $p(\mu'_k|\mu_k)$, and μ_k does not give any additional information to calculate the probability of obtaining c_k when μ'_k is determined. Thus, the likelihood can be expressed as follows.

$$\begin{aligned} p(c_k|\theta_k) &= \int_{-\infty}^{\infty} p(c_k|\mu'_k, \mu_k)p(\mu'_k|\mu_k)\mathrm{d}\mu'_k \\ &= \int_{-\infty}^{\infty} p(c_k|\mu'_k)p(\mu'_k|\mu_k)\mathrm{d}\mu'_k \\ &= \int_{-\infty}^{\infty} \mathcal{N}(c_k - \mu'_k; 0, \sigma_{sen})\mathcal{N}(\mu'_k - \mu_k; 0, \sigma_{env})\mathrm{d}\mu'_k. \end{aligned} \tag{14}$$

Thus, the likelihood can be treated as a convolution of Gaussian distributions, which means it can be expressed as Gaussian distribution again [40], as:

$$\begin{aligned} p(c_k|\theta_k) &= \mathcal{N}(c_k; \mu_k, \sigma_k) \\ &= \frac{1}{\sigma_k\sqrt{2\pi}} \exp\left[-\frac{(c_k - \mu_k)^2}{2\sigma_k^2}\right], \end{aligned} \tag{15}$$

where

$$\sigma_k = \sqrt{\sigma_{k,sen}^2 + \sigma_{k,env}^2} \tag{16}$$

is the standard deviation of the likelihood function. The marginal likelihood $p\left(c_k|c_{1:k-1}\right)$ in Eq. (12) can be calculated using the prior PDF and the new measurement as:

$$p\left(c_k|c_{1:k-1}\right) = \int p\left(c_k|\theta_k\right) p\left(\theta_k|c_{1:k-1}\right) \mathrm{d}\theta_k. \tag{17}$$

B. Particle filter

As the PDF of the source term distribution is nonlinear and non-Gaussian, it is difficult to compute Eq. (12) analytically or exactly. To this end, the particle filter, one of the sampling-based sequential Monte Carlo methods, is used to estimate the PDF of the source term. The particle filter takes samples (called 'particles') to represent the source term including the potential source location and release strength. Each particle, $\theta_k^{(i)}$, is drawn from the proposal (importance) distribution with its associated weight, $w_k^{(i)}$, to approximate the exact source term PDF in Eq. (12) as:

$$p\left(\theta_k|c_{1:k}\right) \approx \sum_{i=1}^{N_p} w_k^{(i)} \delta(\theta - \theta_k^{(i)}), \tag{18}$$

where N_p is the number of the particles and $\delta(\cdot)$ is the Dirac Delta function. As we assume the diffusive source is time-invariant, the sampled particles which represents the estimated source term is not changed by time, that is, $\theta_k^{(i)} = \theta_{k-1}^{(i)}$ for $i = 1, \ldots, N_p$. In this regard, we assume that the proposal distribution for k^{th} step equals the posterior at $(k-1)^{th}$ step as used similarly in [18], [19]. Then, the unnormalized particle weight update based on Bayesian inference for the particle filter can be simply expressed as:

$$w'^{(i)}_k = p(c_k|\theta_k^{(i)}) \cdot w_{k-1}^{(i)}, \tag{19}$$

where $p(c_k|\theta_k^{(i)})$ is calculated by likelihood function in Eq. (15). The normalized weight can be calculated as:

$$w_k^{(i)} = \frac{w'^{(i)}_k}{\sum_{j=1}^{N_p} w'^{(j)}_k}. \tag{20}$$

In general, this type of importance sampling method could cause degeneracy, which means that only a few particles

will have non-zero weights. To overcome the degeneracy, the resampling method is employed with the effective number of samples, n_{eff} [41], given as:

$$n_{eff} \approx \frac{1}{\sum_{i=1}^{N_p}(w_k^{(i)})^2}. \tag{21}$$

Whenever the effective number goes below a certain threshold, the particles are resampled. Additionally, the resampled particles are subjected to the Markov Chain Monte Carlo move step [41] to improve the diversity of the particles; this makes the stationary source term samples $\theta_k^{(i)}$ change its values at each resampling iteration.

IV. Information-Theoretic Source Search

The information-theoretic source search is known to be beneficial in the environment where the gas diffuses sporadically and stochastically such as an atmospheric gas dispersion situation [13], [17]. Infotaxis is one of the popular information-theoretic search methods for source seeking and source term estimation using the information about the source [17], [18], [20]. As it maximizes the reduction of the entropy of the source term at each time step, the search tends to reduce the uncertainty of the source term gradually. The Infotaxis determines the discrete next best action by considering the one-step lookahead (i.e., future) information, so basically it is a greedy algorithm. Here, we propose the algorithm to find the more efficient search path while avoiding the local optima and obstacles by considering multi-step lookahead information. To this end, the receding horizon concept is adopted to plan the multi-step lookahead decision while maintaining the computational loads manageable with the random sampling [42]–[44].

A. Infotaxis

This subsection briefly introduces formulations of original Infotaxis [17] and the modified version of Infotaxis with the particle filter [18]. The Infotaxis is a popular information-theoretic search algorithm that uses the reduction of the entropy for estimation of the source term at each time step with the utility function, $I(u_k)$, as given:

$$\begin{aligned} I(u_k) = &\, p\left(\mathbf{p}_{k+1}\right)\left[H_k\right] \\ &- \left(1 - p\left(\mathbf{p}_{k+1}\right)\right)\left[E_{\hat{c}}\left[\hat{H}_{k+1}\left(u_k\right)\right] - H_k\right], \end{aligned} \tag{22}$$

where

$$H_k = -\int p\left(\theta_k|c_{1:k}\right)\log p\left(\theta_k|c_{1:k}\right)\mathsf{d}\theta_k. \tag{23}$$

The probability, $p(\mathbf{p}_{k+1})$, represents the estimated probability of the source occurring at the next sensing position, $\mathbf{p}_{k+1}$ with a choice of the control action $u_k \in \mathbf{u} = [\uparrow, \downarrow, \leftarrow, \rightarrow]$. The Shannon's entropy for a source term is represented as H_k and $E_{\hat{c}}[\cdot]$ means the expectation over the random variable $\hat{c}$ (i.e., predicted measurements) which is determined by control decision u_k. The current optimal decision, u_k^*, is selected to maximize the utility function, $I(u_k)$. In this study, we use the simplified version of the utility function assuming that $p\left(\mathbf{p}_{k+1}\right) = 0$, termed as Infotaxis II, to reduce computational burden [18] as:

$$I(u_k) = H_k - E_{\hat{c}}\left[\hat{H}_{k+1}\left(u_k\right)\right]. \tag{24}$$

The particle filter makes Eq. (24) easy to be calculated. According to Eq. (18), as the posterior probability, $p\left(\theta_k|c_{1:k}\right)$, can be represented by the weight of the particle filter, the approximation of the Shannon's entropy using the particle filter is expressed as:

$$H_k \approx -\sum_{i=1}^{N_p} w_k^{(i)} \log w_k^{(i)}, \tag{25}$$

where N_p is the number of particles (potential source terms) of the particle filter. The expected Shannon's entropy using the possible future measurements $\hat{c}_{k+1}$ at $\mathbf{p}_{k+1}$, determined by the current sensing position $\mathbf{p}_k$ and control decision u_k, is calculated as:

$$E_{\hat{c}}\left[\hat{H}_{k+1}\left(u_k\right)\right] = \int p\left(\hat{c}_{k+1}(u_k)|\theta_k\right)\hat{H}_{k+1}\left(u_k\right)\mathsf{d}\hat{c}_{k+1}, \tag{26}$$

where

$$\hat{H}_{k+1}\left(u_k\right) = -\int \log p(\theta_k|c_{1:k}, \hat{c}_{k+1}(u_k))\mathsf{d}\theta_k. \tag{27}$$

The expected PDF of the source term, $p(\theta_k|c_{1:k}, \hat{c}_{k+1})$, can also be approximated by the particle filter. The unnormalized weight of the potential source terms can be updated similar to Eq. (19):

$$\hat{w}_{k+1}^{\prime(i)} = p\left(\hat{c}_{k+1}(u_k)|\theta_k^{(i)}\right)\cdot w_k^{(i)}, \tag{28}$$

where the normalized weight of $\hat{w}_{k+1}^{\prime(i)}$, i.e, $\hat{w}_{k+1}^{(i)}$ can be calculated similar to Eq. (20). The expectation of the entropy at $(k+1)^{th}$ time step can then be re-expressed as:

$$\begin{aligned} E_{\hat{c}}\left[\hat{H}_{k+1}\left(u_k\right)\right] \approx -\int &\, p\left(\hat{c}_{k+1}(u_k)|\theta_k\right) \\ &\cdot \sum_{i=1}^{N_p} \hat{w}_{k+1}^{(i)} \log \hat{w}_{k+1}^{(i)} \mathsf{d}\hat{c}_{k+1}, \end{aligned} \tag{29}$$

By substituting Eq. (23) and (29) into Eq. (24), the approximated utility function with the particle filter can be obtained:

$$\begin{aligned} I(u_k) =& H_k - E_{\hat{c}}\left[\hat{H}_{k+1}\left(u_k\right)\right] \\ \approx& -\sum_{i=1}^{N_p} w_k^{(i)} \log w_k^{(i)} \\ &+ \int p\left(\hat{c}_{k+1}(u_k)|\theta_k\right)\cdot \sum_{j=1}^{N_p} \hat{w}_{k+1}^{(j)} \log \hat{w}_{k+1}^{(j)} \mathsf{d}\hat{c}_{k+1}. \end{aligned} \tag{30}$$

Since the posterior source term distribution approximated by the particle filter is highly non-linear and non-Gaussian, the second term on the right hand side in Eq. (30) cannot be analytically calculated. Thus, in this study, instead of the continuous future measurement $\hat{c}_{k+1}$, the discretized measurement set with the certain discretization interval, $\delta\hat{d}$, is used as:

$$\begin{aligned}\hat{\mathbf{d}}_{k+1} &= [\min(\hat{d}_{k+1}), \min(\hat{d}_{k+1}) + \delta\hat{d}_{k+1}, \cdots, \max(\hat{d}_{k+1})] \\ &= [{}^{(1)}\hat{d}_{k+1}, {}^{(2)}\hat{d}_{k+1}, \cdots, {}^{(N_d)}\hat{d}_{k+1}]\end{aligned} \tag{31}$$

where $\min(\hat{d}_{k+1}) = {}^{(1)}\hat{d}_{k+1}$ and $\max(\hat{d}_{k+1}) = {}^{(N_d)}\hat{d}_{k+1}(= \min(\hat{d}_{k+1}) + (N_d - 1)\delta\hat{d}_{k+1})$ represent the possible minimum and maximum measurements at $(k+1)^{th}$ time step, respectively, and the number of discretized measurements is represented as N_d. We set minimum and maximum possible measurement values using the empirical three-sigma rule as $\mu_{k+1} \pm 3 \cdot \sigma_{k+1}$. Here, $\mu_{k+1} = \sum_{i=1}^{N_p} \mu(\mathbf{p}_{k+1}; \theta_k^{(i)}) w_k^{(i)}$ represents the expected mean concentration calculated by Eq. (1) and $\sigma_{k+1}^{(i)} = \sqrt{(\sigma_{env})^2 + (\sigma_{k+1,sen}^{(i)})^2}$ as in Eq. (16). The utility function using the discretized Gaussian sensor model can then be re-written as:

$$\begin{aligned}I(u_k) =& H_k - E_{\hat{d}}\left[\hat{H}_{k+1}(u_k)\right] \\ \approx& -\sum_{i=1}^{N_p} w_k^{(i)} \log w_k^{(i)} \\ &+ \sum_{l=1}^{N_d} p({}^{(l)}\hat{d}_{k+1}|\theta_k) \cdot \sum_{j=1}^{N_p} \hat{w}_{k+1}^{(j)} \log \hat{w}_{k+1}^{(j)}.\end{aligned} \tag{32}$$

Since we use the Gaussian sensor model, the shape of the measurement probability distribution from each potential source term (i.e., particle) follows the Gaussian distribution. Thus, the discretized conditional probability, $p({}^{(l)}d_{k+1}|\theta_k)$, can be calculated by summation of the cumulative distribution function (CDF) of the standard normal distribution, $\Phi(\cdot)$ [37], with the associated particle weight, $w_k^{(i)}$, as:

$$\begin{aligned}&p({}^{(l)}\hat{d}_{k+1}|\theta_k) \\ &= \sum_{i=1}^{N_p} \left[\Phi\left(\frac{\Delta^{(l+1)}\hat{d}_{k+1}}{\sigma_{t,k+1}^{(i)}}\right) - \Phi\left(\frac{\Delta^{(l)}\hat{d}_{k+1}}{\sigma_{t,k+1}^{(i)}}\right)\right] w_k^{(i)},\end{aligned} \tag{33}$$

where

$$\begin{aligned}\Delta^{(l)}\hat{d}_{k+1} &= {}^{(l)}\hat{d}_{k+1} - \mu(\mathbf{p}_{k+1}; \theta_k^{(i)}), \\ \sigma_{k+1}^{(i)} &= \sqrt{(\sigma_{env})^2 + (\sigma_{k+1,sen}^{(i)})^2}.\end{aligned} \tag{34}$$

The Infotaxis strategy guides the mobile sensor to obtain a measurement in a direction which maximizes the reduction of the expected entropy of the potential source term PDF. Therefore, the optimal decision $u_k^* \in \mathbf{u} = [\uparrow, \downarrow, \rightarrow, \leftarrow]$ is selected to maximize $I(u_k)$ as:

$$u_k^* = \underset{u_k \in \mathbf{u}}{\operatorname{argmax}}\, I(u_k). \tag{35}$$

It is worthwhile noting that the original Infotaxis is a greedy approach that uses one-step lookahead information (i.e., $\hat{c}_{k+1}$ at $\mathbf{p}_{k+1}$) only, so we call it greedy Infotaxis hereafter to avoid confusion between the greedy Infotaxis and proposed receding horizon Infotaxis. The simulation result for an illustrative run using the greedy Infotaxis in an obstacle-rich environment is shown in Fig. 3. The mobile sensor is equipped with a LiDAR with a radius of 24 meters to create a local map of an unknown obstacle environment. This obstacle map is used to generate available paths for source term estimation. The search trajectory and particles are shown in the lower graph. The blue star means the current mobile sensor position, the blue line is the search trajectory, and the blue dots are particles from the particle filter. The red dots represent the measurement where the bigger size means higher concentration. The gray environment is an unknown area and the visited area is represented with black and white.

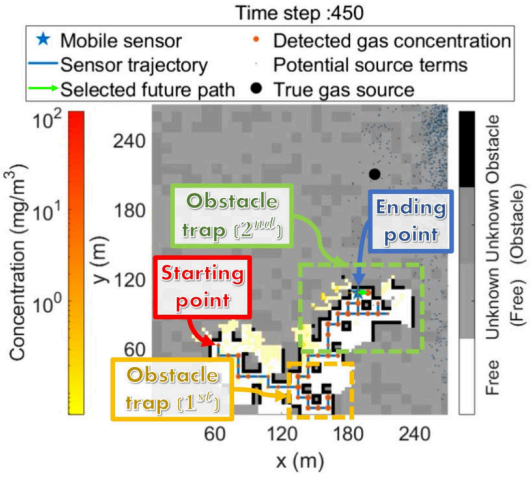


Fig. 3. The example path of the greedy Infotaxis in an urban environment.

As shown in Fig. 3, the mobile sensor can be easily trapped in a space surrounded by obstacles (represented as obstacle trap) as it considers one-step lookahead information/control action only, which shows the limitation of a greedy-based algorithm in the complex environment. For instance, when the mobile sensor encounters an obstacle and the estimated source location happens to be at the other side of the obstacle, it first moves one step towards the obstacle; it is expected to have better information getting closer to the source location. At the next step, it realizes that it cannot move any further to the same direction due to the obstacle, so it chooses the second best direction. However, at the next step, it attempts to move towards the obstacle again to gain better information and this behavior of wandering around the obstacle is repeated. By the random nature of the greedy Infotaxis from the particle filter, it might escape the obstacle after a while. To address this obstacle trap issue, receding horizon-based Infotaxis is introduced in the following section.

B. Receding Horizon Infotaxis (RHI)

In this section, we propose the receding horizon Infotaxis (RHI) strategy using the Gaussian sensor model and binary sensor model. A receding horizon concept is used to determine a multi-step lookahead search path as illustrated in Fig. 4. This method calculates the optimal decision sequence from the utility function of multiple steps for a certain duration of time but only takes the first control decision recursively. To avoid generating inefficient search path, action sequence that return to the original position after two steps are excluded from consideration. The receding horizon method becomes computationally expensive as the number of horizon steps increases, and the computational complexity also depends on the

dimension of future states to be considered. We consider two sensor models to predict the future states: the Gaussian sensor model and the binary sensor model. The Gaussian sensor model is based on the white Gaussian noise. It can represent different measurement values densely, but at the same time, it increases the computational burden when predicting possible future measurements. The binary sensor model divides the measurement into only two values: the detection or the non-detection based on the sensing threshold, and thus it has less computational complexity compared with the Gaussian sensor model. The random sampling method is also used to further reduce the computational complexity. This method considers only a few sampled future measurement sequences from the estimated source probability distribution rather than considering all possible future measurement sequences in the receding horizon steps. From the following subsections, several proposed RHI methods will explained in detail.

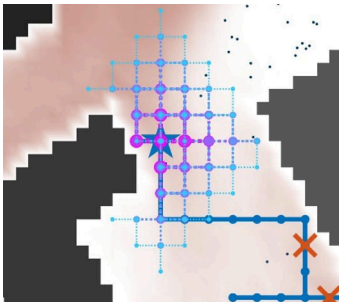


Fig. 4. An illustrative search path of multi-step lookahead decision making. The thick blue line represents the search path, the black grids represent buildings, and the blue star indicates the current position of the mobile sensor. The dashed lines are possible decision paths in multiple steps where a deeper future step corresponds to a thinner line.

1) ***Receding horizon Infotaxis with Gaussian sensor model (RHI-G):*** The RHI using the Gaussian sensor model is called RHI-G in this study. The mobile sensor selects the optimal decision sequence for a certain time horizon length K from the current time step k, $u^*_{k:k+K-1} = \left[u^*_k, u^*_{k+1}, \cdots, u^*_{k+K-1}\right]$, among all possible decision sequences, $\mathbf{U}_{k:k+K-1} = \left[u^{(path)}_{k:k+K-1}\right]_{1\le path\le N_u^K}$ where $u^{(path)}_{k:k+K-1} = \left[u^{(path)}_k, u^{(path)}_{k+1}, \cdots, u^{(path)}_{k+K-1}\right]$ is a $(path)^{th}$ sample decision sequence and each decision (control action) satisfies $u^{(path)} \in U = [\uparrow, \downarrow, \leftarrow, \rightarrow]$. The total number of possible decision sets is equal to $N_u^K = 4^K$. The computation of the utility function using the $(path)^{th}$ decision sequence is given as:

$$I(u^{(path)}_{k:k+K-1}) = \sum_{n=1}^{K} r^n E_{\hat{\mathbf{D}}}\left[I^{(path,m)}_{k+n-1}\right], \tag{36}$$

where m is used to represent m^{th} future measurement sample which can be obtained along the $(path)^{th}$ decision sequence sample. Besides, the belief ratio $0 \le r \le 1$ is introduced as a discount factor, which means we less believe the further predicted utility function. The expectation of the predicted utility function for $(k+n-1)^{th}$ time step is represented as $E_{\hat{\mathbf{D}}}\left[I^{(path,m)}_{k+n-1}\right]$, which is computed by the future measurement permutation set $\hat{\mathbf{D}} = \hat{\mathbf{D}}_{k+1:k+n-1} = \left[\hat{d}^{(m)}_{k+1:k+n-1}\right]_{1\le m\le N_d^{n-1}}$ where N_d represents the number of possible discretized measurements as mentioned in Eq. (31). Similar to the possible decision sequence, a sample measurement sequence is expressed as $\hat{d}^{(m)}_{k+1:k+n-1} = \left[\hat{d}^{(m)}_{k+1}, \hat{d}^{(m)}_{k+2}, \cdots, \hat{d}^{(m)}_{k+n-1}\right]$ for the given $(path)^{th}$ decision sequence as illustrated by the red dashed line in Fig. 5. Note that it is assumed $\hat{d}^{(m)}_{k+1:k} = \emptyset$ when $n = 1$. Each future measurement can be obtained from the discretized measurement set using Eq. (31) where the minimum and maximum measurements are determined by the expected mean concentration at the corresponding sensing position with the estimated source term as illustrated in Fig. 5.

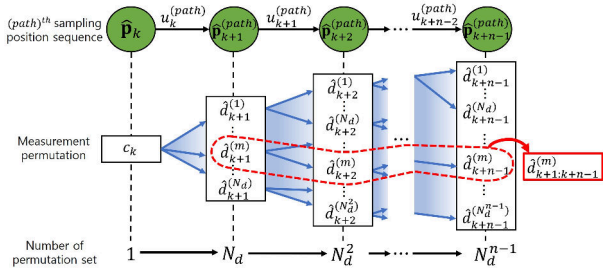


Fig. 5. The measurement permutation sets for the $(path)^{th}$ decision sequence and corresponding sensing position sequence.

For each selected decision sequence, the predicted utility $I^{(path,m)}_{k+n-1}$ is the function of the random variable vector $\hat{d}^{(m)}_{k+1:k+n-1}$. Thus, the expectation of the RHI-G utility function component at $(k+n-1)^{th}$ time step in the $(path)^{th}$ sampled decision sequence can be represented as:

$$E_{\hat{\mathbf{D}}}\left[I^{(path,m)}_{k+n-1}\right] = \sum_{m=1}^{N_d^{n-1}} p(\hat{d}^{(m)}_{k+1:k+n-1}|c_{1:k}) I^{(path,m)}_{k+n-1}, \tag{37}$$

where

$$\begin{aligned} p(\hat{d}^{(m)}_{k+1:k+n-1}|c_{1:k}) &= p(\hat{d}^{(m)}_{k+1}, \cdots, \hat{d}^{(m)}_{k+n-1}|c_{1:k}) \\ &= p(\hat{d}^{(m)}_{k+n-1}|\hat{d}^{(m)}_{k+n-2}, \cdots, c_{1:k}) \\ &\quad \cdot p(\hat{d}^{(m)}_{k+n-2}|\hat{d}^{(m)}_{k+n-3}, \cdots, c_{1:k}) \cdots p(c_{1:k}) \\ &= \prod_{a=1}^{n-1} p(\hat{d}^{(m)}_{k+a}|\hat{d}^{(m)}_{k+1:k+a-1}, c_{1:k}). \end{aligned} \tag{38}$$

Note that $\sum_{m=1}^{N_d} p(\hat{d}^{(m)}_{k+1:k+n-1}) = 1$ for a given $(path)^{th}$ decision sequence. Each conditional distribution in Eq. (38) is the marginal likelihood distribution which can be calculated similar to Eq. (17). As we use the particle filter to estimate the source term, each conditional probability can be calculated as:

$$\begin{aligned} p(\hat{d}^{(m)}_{k+a}|\hat{d}^{(m)}_{k+1:k+a-1}, c_{1:k}) &= \sum_{i=1}^{N_p} p(\hat{d}^{(m)}_{k+a}|\theta^{(i)}_k) \\ &\quad \cdot p(\theta^{(i)}_k|\hat{d}^{(m)}_{k+1:k+a-1}, c_{1:k}) \\ &= \sum_{i=1}^{N_p} p(\hat{d}^{(m)}_{k+a}|\theta^{(i)}_k) \hat{w}^{(i)}_{k+a-1}. \end{aligned} \tag{39}$$

The likelihood function $p(\hat{d}^{(m)}_{k+a}|\theta^{(i)}_k)$ can be calculated using Eq. (33) and the estimated probability of the source term. The estimated source term probability is expressed in the particle filter weight, $\hat{w}^{(i)}_{k+a-1}$, that is updated using Bayesian inference as in Eq. (28) as:

$$\begin{aligned} p(\theta^{(i)}_k|\hat{d}^{(m)}_{k+1:k+a-1}, c_{1:k}) &= \hat{w}^{(i)}_{k+a-1} \\ &\propto p(\hat{d}^{(m)}_{k+a-1}|\theta^{(i)}_k)\hat{w}^{(i)}_{k+a-2} \\ &\propto \Pi^{a-1}_{\alpha=1} p(\hat{d}^{(m)}_{k+\alpha}|\theta^{(i)}_k) w^{(i)}_k. \end{aligned} \tag{40}$$

The component of the RHI-G utility function for a given $(path)^{th}$ decision sequence and m^{th} measurement sequence in Eq. (37) is calculated as:

$$\begin{aligned} I^{(path,m)}_{k+n-1} =& \hat{H}^{(path,m)}_{k+n-1} - E_{\hat{d}^{(m)}_{k+n}}[\hat{H}^{(path,m)}_{k+n}] \\ =& -\sum_{i=1}^{N_p} \hat{w}^{(i)}_{k+n-1} \log \hat{w}^{(i)}_{k+n-1} \\ &+ \sum_{l=1}^{N_d} p({}^{(l)}\hat{d}_{k+n}|\theta_k) \sum_{j=1}^{N_p} \hat{w}^{(j)}_{k+n} \log \hat{w}^{(j)}_{k+n}, \end{aligned} \tag{41}$$

where it is the utility function at $(k+n-1)^{th}$ time step as in Eq. (32), and $\hat{w}^{(i)}_{k+n-1}$ represents the approximated weight of the potential source term which can be sequentially updated similar to Eq. (28) with a given m^{th} predicted future measurement permutation $\hat{d}^{(m)}_{k+1:k+n-1}$ for a given future decision $\hat{u}^{(path)}_{k-1:k+n-2}$. The weight $\hat{w}^{(i)}_{k+n}$ is updated using the future measurement ${}^{(l)}\hat{d}_{k+n}$ of $(k+n)^{th}$ time step.

The computational cost of this method increases exponentially with the length of the time horizon as this method requires calculating the probability of all possible future measurement sequences and feasible decision (control action) sequences. The computational complexity of calculating the utility function increased by the time horizon K can be expressed as:

$$\mathcal{O}(N_d N_p \cdot N_d^{K-1} \cdot N_u^K), \tag{42}$$

where N_d is the number of discretized measurements, N_p is the number of particles, and N_u is the number of feasible decisions (control actions) for each time step. The computational complexity for calculating the component of the RHI-G utility function $I^{(path,m)}_{k+n-1}$, in Eq. (41) is equal to $N_d N_p$. The total number of future measurement sequences is N_d^{K-1} and the total number of possible decision sequences is N_u^K. It is necessary to reduce the computational complexity to implement the multi-step lookahead decision making in real-time, which will be explained in the following sections.

*2) **Receding horizon Infotaxis with binary sensor model (RHI-B)**:* The receding horizon Infotaxis with binary sensor model (RHI-B) uses the binary sensor model to predict the future measurement sequences. The binary sensor model has only two possible sensing values: 1 (above threshold) or 0 (below threshold). This can reduce the total number of possible measurement sequences significantly compared with the Gaussian sensor model. Similar to the utility function of the RHI-G, the utility function of the RHI-B for a given $(path)^{th}$ decision sequence using the binary sensor model can be expressed as:

$$\begin{aligned} I(u^{(path)}_{k:k+K-1}) &= \sum_{n=1}^{K} r^n E_{\hat{\mathbf{B}}}\left[I^{(path,m)}_{k+n-1}\right] \\ &= \sum_{n=1}^{K} r^n \sum_{m=1}^{N_d^{n-1}} p(\hat{\mathbf{b}}^{(m)}_{k+1:k+n-1}|c_{1:k}) I^{(path,m)}_{k+n-1}, \end{aligned} \tag{43}$$

where

$$p(\hat{\mathbf{b}}^{(m)}_{k+1:k+n-1}|c_{1:k}) = \prod_{a=1}^{n-1} \sum_{i=1}^{N_p} p(\hat{b}^{(m)}_{k+a}|\theta^{(i)}_k)\hat{w}^{(i)}_{k+a-1} \tag{44}$$

$$\begin{aligned} I^{(path,m)}_{k+n-1} =& -\sum_{i=1}^{N_p} \hat{w}^{(i)}_{k+n-1} \log \hat{w}^{(i)}_{k+n-1} \\ &+ \sum_{\hat{b}_{k+n}=0}^{1} p(\hat{b}_{k+n}|\theta_k) \sum_{j=1}^{N_p} \hat{w}^{(j)}_{k+n} \log \hat{w}^{(j)}_{k+n}. \end{aligned} \tag{45}$$

The binary future measurement permutation set $\hat{\mathbf{B}}$ is given as:

$$\hat{\mathbf{B}} = \hat{\mathbf{B}}_{k+1:k+n-1} = \left[\hat{\mathbf{b}}^{(m)}_{k+1:k+n-1}\right]_{1\le m\le 2^{n-1}}, \tag{46}$$

where

$$\hat{\mathbf{b}}^{(m)}_{k+1:k+n-1} = \left[\hat{b}^{(m)}_{k+1}, \hat{b}^{(m)}_{k+2}, \cdots, \hat{b}^{(m)}_{k+n-1}\right]. \tag{47}$$

The number of possible measurements for each time step is two (i.e., $N_d = 2$) for RHI-B. Although the computational complexity of this method is much lower than that of RHI-G, computational complexity still increases exponentially with respect to the length of the time horizon as:

$$\mathcal{O}(2N_p \cdot 2^{K-1} \cdot N_u^K). \tag{48}$$

*3) **Receding horizon Infotaxis with random sampling (RHI-R)**:* In order to further reduce the computational complexity of aforementioned approaches, we introduce the RHI algorithm with the random sampling method (RHI-R) that samples the future measurement sequence using the sensor model and prior estimated source term probability distribution. The RHI-R includes RHI-GR and RHI-BR which follow the same methodology but using Gaussian and binary sensor models, respectively. Here, explanations are based on RHI-BR. Each component in the predicted future measurement sequence at $(k+n)^{th}$ time step, $\hat{b}^{(m)}_{k+n}$, is drawn from the estimated source term distribution, $p(\hat{b}^{(m)}_{k+n}|c_{1:k}, \hat{\mathbf{b}}^{(m)}_{k+1:k+n-1})$, which is approximated by the particle filter as:

$$\hat{b}^{(m)}_{k+n} \sim p(\hat{b}^{(m)}_{k+n}|c_{1:k}, \hat{\mathbf{b}}^{(m)}_{k+1:k+n-1}) = \sum_{i=1}^{N_p} p(\hat{b}^{(m)}_{k+n}|\theta^{(i)}_k)\hat{w}^{(i)}_{k+n-1}, \tag{49}$$

where $p(\hat{b}^{(m)}_{k+n}|\theta^{(i)}_k)$ can be calculated using Eq. (7). The example of random sampling is shown in Fig. 6.

Previously, the utility function of the RHI is computed while taking into account all predicted future measurement permutation sequences. However, in RHI-BR, measurements

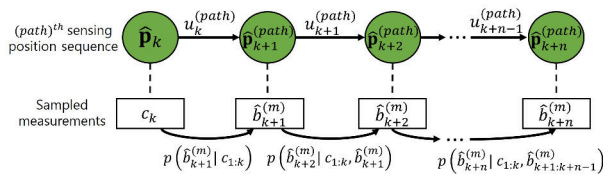


Fig. 6. The example of the sampled measurement sequence for the $(path)^{th}$ decision sequence.

are randomly sampled using the estimated measurement distribution in each $(k+n)^{th}$ time step in Eq. (49). The utility function is calculated using the sampled future measurement sequence. The utility function is no longer a summation of expectation as the sampled measurements are no longer a random variable. Thus, the utility function is calculated as:

$$I(u_{k:k+K-1}^{(path)}) = \sum_{n=1}^{K} r^n \sum_{m=1}^{N_m} \frac{I_{k+n-1}^{(path,m)}}{N_m}, \tag{50}$$

where $I_{k+n-1}^{(path,m)}$ can be calculated by using Eq. (45). With the random sampling approximation, the total number of measurement sequences m is not N_d^K but N_m. By applying this sampling approach, the computational complexity no longer increases with the number of possible future measurements sequences along the length of the time horizon; it depends on the number of sampled measurement sequences, N_m, which is a design parameter. The computational complexity of this method for the worst case is then reduced as:

$$\mathcal{O}(2N_p \cdot N_m \cdot N_u^K). \tag{51}$$

Similarly, the computational complexity for RHI-GR is given as:

$$\mathcal{O}(N_d N_p \cdot N_m \cdot N_u^K). \tag{52}$$

The pseudo-code of the RHI-BR at k^{th} time step is given in Algorithm 1. Note that, since the mobile agent has only partial information until the end of the search and the gas sensor measurement is highly stochastic, the future prediction is inaccurate. Thus, we use only the first element in the optimal decision sequence, u_k^*, for actual movement. Besides, when a certain path from a decision sequence intersects with obstacles, it is discarded as represented in line 8. The obstacle map is recursively updated by using the local range sensor (e.g., LiDAR) observation, represented by the occupancy grid map [45] as previously shown in Fig. 3.

Algorithm 1 RHI-BR algorithm

1: $c_k \leftarrow$ obtain a new sensor measurement
2: $Obs_k \leftarrow Obs_{k-1}$ ▷ Update obstacle map
3: $\{(\theta_{k-1}, w_{k-1})\} \rightarrow \{(\theta_k, w_k)\}$ ▷ Update using Eq. (19)
4: **for** $path = 1, \cdots, N_u^K$ **do** ▷ For all decision sequences
5: **for** $m = 1, \cdots, N_m$ **do** ▷ For all sampled measurement sequences
6: **for** $n = 1, \cdots, K$ **do**
7: $\hat{\mathbf{p}}_{k+n}^{(path)} = \mathbf{p}_k + u_{k+n-1}^{(path)}$ ▷ Update predicted sensing positions
8: **if** $Obs_k(\hat{\mathbf{p}}_{k+n}^{(path)}) == 0$ **then** ▷ If no obstructed area
9: $\hat{b}_{k+1:k+n-1}^{(m)}$ ▷ Sample measurements using Eq. (49)
10: $I_{k+n-1}^{(path,m)}$ ▷ Calculate using Eq. (45)
11: **else**
12: break
13: $I(u_{k:k+K-1}^{(path)}) = \sum_n r^n \sum_m I_{k+n-1}^{(path,m)}/N_m$ ▷ Calculate using Eq. (50)
14: $u_{k:k+K-1}^* = \underset{u^{(path)} \in \mathbf{U}_{k:k+K-1}}{\operatorname{argmax}} I(u_{k:k+K-1}^{(path)})$ ▷ Choose the best action
15: $\mathbf{p}_{k+1} = [x_{k+1}\ y_{k+1}]^\mathsf{T} = [x_k\ y_k]^\mathsf{T} + u_k^*$ ▷ Move to a new sensing position

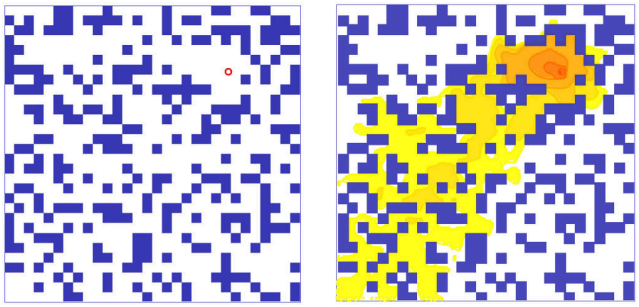
V. Numerical Simulations

This section provides illustrative runs of source search and estimation using the proposed receding horizon Infotaxis (RHI) methods including RHI-G, RHI-B, RHI-GR, and RHI-BR. To provide comparisons and show the advantage of the proposed approaches, three sets of experiments are discussed. First, the search performance of the proposed methods are compared in Section V-B. The search performance is compared in terms of the success rate, the mean search time, and the computational time. When the standard deviation of the estimated source location is lower than a certain threshold (3m) and the mean value of the estimation is close enough to the ground truth (3m) before reaching the maximum time step (150 steps), the search is considered as a success. The mean search time is averaged search time steps of successful cases and the success rate means the ratio of the successful cases over the Monte Carlo simulations. The computational time is the time taken for each algorithm to decide the next control decision per time step. It is assumed that the sensing and moving time between grid points for each time step of the mobile sensor are same for all simulations. Then, to show the advantage of using the receding horizon-based approach more clearly, the result with a specific scenario is discussed in details in Section V-C. The proposed search strategies are compared with the state-of-the-art source search strategy developed for obstacle-rich environments, named IWFA-2 [30], on a realistic scenario in Section V-D. Lastly, the hardware-in-the-loop simulation in 3D environment is presented in Section V-E.

A. Simulation environments

This section describes the hazardous substance diffusion scenarios in randomly generated obstacle-rich environments. To consider the general shape of obstacles in a two-dimensional space, we assumed obstacles consist of grids, where the size of each grid is $9m \times 9m$. The various obstacles are shown in blue grids, white grids are free space, and red circles represent gas emission source in Fig. 7. In order to generate connected free space components from a starting point to the source location by randomly generated obstacles, the probability of the presence of obstacle ρ_0 for each grid must satisfy $\rho_0 < 1 - \rho_c$, where ρ_c means the site percolation threshold [46], [47]. The ρ_c for the square grid is set to 0.593.

The Graz Lagrangian (GRAL) model is adopted to simulate the gas dispersion in the urban environment [48]. This model is based on the Lagrangian particle dispersion model which is one of the computational fluid dynamic (CFD) models. This model calculates a trace of gas particles from the diffusive source by considering gas properties and statistical environments. The three-dimensional wind model is obtained using the Reynolds-averaged Navier-Stokes equations. The effect of buildings (i.e., obstacles) is calculated by using the standard $k-\epsilon$ turbulence model [48]. An example gas dispersion map based on the obstacle map presented in Fig. 7(a) is shown in Fig. 7(b). Note that, this GRAL model is quite realistic and different from the ideal gas dispersion model (Eq. (1)) used in the estimation algorithm.



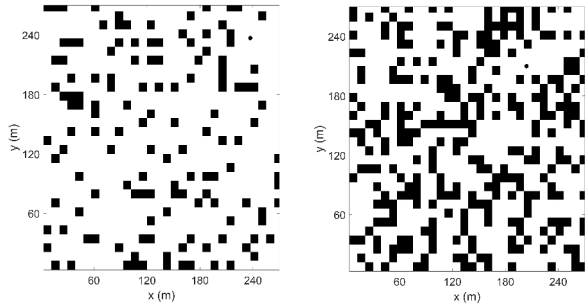
(a) The example of obstacle map

(b) The example of gas dispersion

Fig. 7. The example obstacle-rich environment and corresponding gas dispersion result using the realistic CFD model.

B. Performance comparison for various RHI methods

We examined the search performance according to the horizon length and approximation methods of RHI in terms of search time and success rate for the randomly generated obstacle maps with the different obstacle generation probabilities $\rho_0 = 0$, 0.15, and 0.3. It can be considered that $\rho_0 = 0.3$ represents an obstacle-rich environment as shown in Fig. 8(a), $\rho_0 = 0.15$ an environment with relatively few obstacles as shown in Fig. 8(b), and $\rho_0 = 0$ no obstacles.



(a) Obstacle map ($\rho_0 = 0.15$)

(b) Obstacle map ($\rho_0 = 0.3$)

Fig. 8. The sample maps for obstacle generation probability of $\rho_0 = 0.15$ and 0.3.

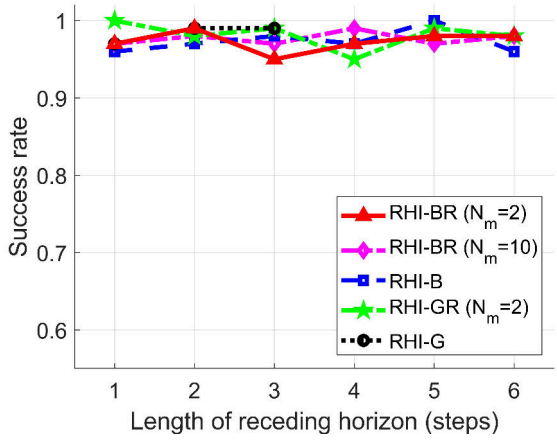
Monte Carlo simulation results on those maps are shown in Fig. 9. As the computational time of RHI-G increases significantly after three steps of the horizon length (as will be displayed in Fig. 10), it is limited to three steps. For the fair comparison of all methods, a hundred different randomly-generated obstacle maps are used for each ρ_0 while the initial location of the mobile sensor is fixed at $(60, 60)$. To analyze the source search performance influenced by decision making when easy to estimate the source, the true gas dispersion model for these Monte Carlo simulations is the same as the one for estimation as in Eq. (1), that is, the GRAL model is not used here. The simulation parameters are set as:

- Search area $A = 270m \times 270m$, wind direction $\phi = 220°$ with velocity $V = 2m/s$, the effective diffusivity parameter $\zeta = 10n^2/s$, release strength of the substance $Q = 7.2kg/h = 2000mg/s$, particle lifetime $\tau = 1000s$, and source location $\mathbf{p}_s = [235m, 235m, 12m]^\mathsf{T}$; and
- Feasible control actions, $U = [\uparrow, \downarrow, \leftarrow, \rightarrow]$, and the moving distance of each step $9m$, the standard deviation of environmental and sensor noise are $\sigma_{env} = 0.4mg/m^3$ and $\sigma_{sen} = 0.2 \cdot \mu(p;\theta)mg/m^3$ respectively, the number of particles $N_p = 1,000$ for the particle filter, the belief ratio $r = 0.8$, binary measurement threshold update parameter $\lambda = 0.8$, and the number of sampled future measurement sequences $N_m = 2$.

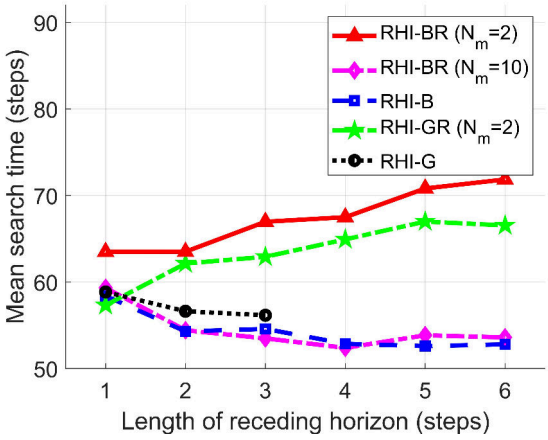
Note that, the performance of the one horizon step in Fig. 9 indicates the search performance of the greedy Infotaxis. As shown in Fig. 9(a) and (c), there is no significant difference in success rate for $\rho_0 = 0$ and 0.15 cases, and even using the greedy Infotaxis, the source term can be estimated with a probability of more than 90%. However, in the obstacle-rich environment ($\rho_0 = 0.3$), the success rate of source term estimation using the greedy Infotaxis is only about 60%, while success rates of proposed algorithms using the receding horizon concept increase as the length of horizon increases. The success rate of the proposed algorithms is about 90% when it uses six horizon steps.

The proposed RHI-G and RHI-B can find the source with fewer mean search time steps than that of the greedy Infotaxis in no-obstacle environments. The performance of algorithms using the sampling method (RHI-GR and RHI-BR) with a few sampled future measurement sequences ($N_m = 2$) gets even worse as the horizon increases as shown in Fig. 9(b) whereas RHI-BR with the enough number of samples ($N_m = 10$) shows improvements with longer horizon steps. As the obstacle ratio increases, the mean search time for all cases is significantly reduced as the horizon length increases as shown in Fig. 9(d) and (f). RHI-B shows the best performance, followed by RHI-GR and RHI-BR when it uses six horizon steps. Since the search performances of RHI-BR and RHI-GR are good enough with a few sampled measurement sequences ($N_m = 2$) compared with greedy Infotaxis, the large number of measurement sampling case ($N_m = 10$) is not considered in obstacle environments.

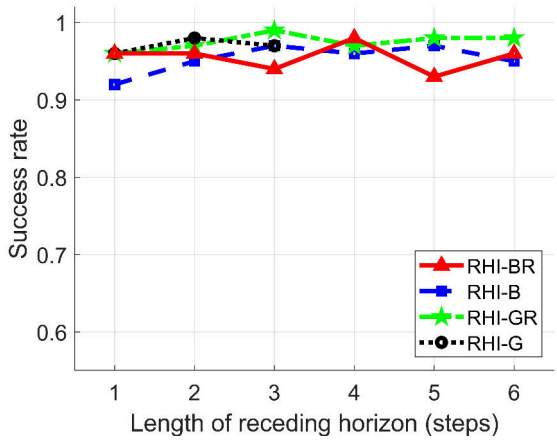
The computation time using MATLAB on a desktop computer with an Intel(R) Core(TM) i7 7700 CPU @ 3.60 GHz is compared to check the computational efficiency of the proposed algorithms. The computation time of each algorithm



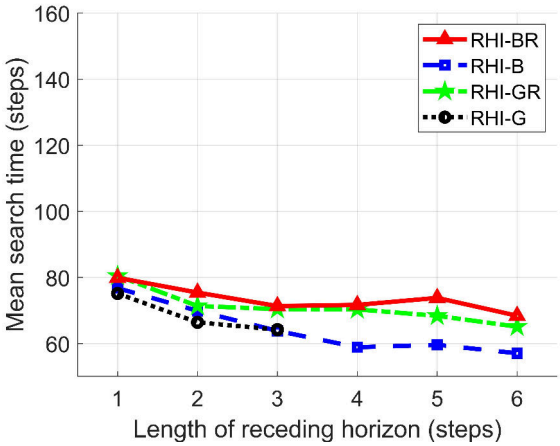
(a) Success rate ($\rho_0 = 0$)



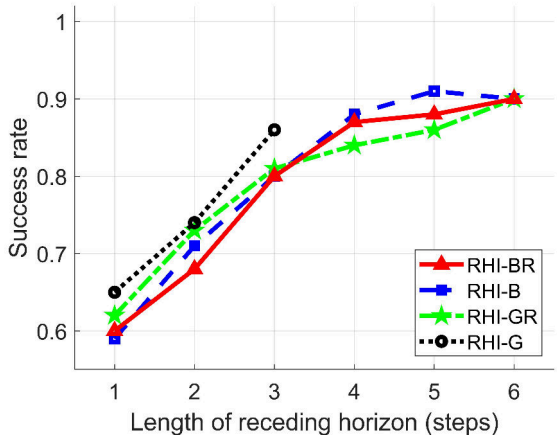
(b) Mean search time ($\rho_0 = 0$)



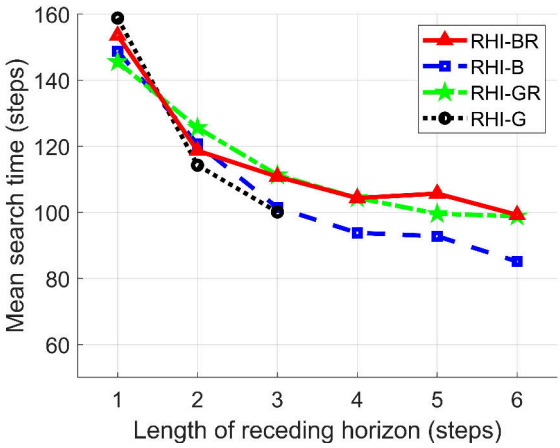
(c) Success rate ($\rho_0 = 0.15$)



(d) Mean search time ($\rho_0 = 0.15$)



(e) Success rate ($\rho_0 = 0.3$)



(f) Mean search time ($\rho_0 = 0.3$)

Fig. 9. The results of Monte Carlo simulation with various approximation methods and horizon length under obstacle maps.

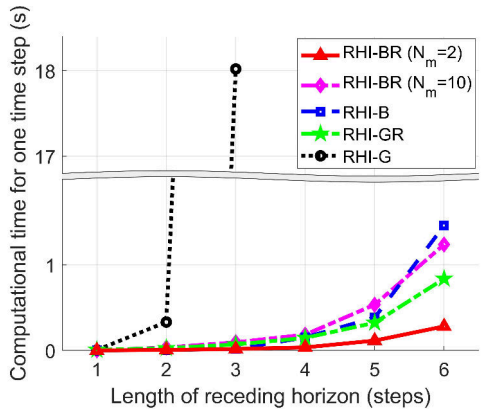


Fig. 10. The results of Monte Carlo simulation with various approximation methods and horizon length under $\rho_0 = 0.3$ obstacle probability map.

per one time step according to the length of the receding horizon is shown in Fig. 10. As explained in Section IV-B, RHI-G has the highest computational burden, followed by RHI-B and then the random sampling-based approaches. RHI-BR ($N_m = 2$) has the lowest computational cost than as expected.

From above Monte Carlo simulation results, it can be concluded that as the horizon length increases, the success rate increases in the obstacle-rich environment, but there is no significant difference among proposed algorithms. The mean search time is decreased as the horizon length increases in the environment even with a few or no obstacles, and there are some differences between algorithms. The random sampling method can decrease the computational burden at the cost of the increased search time steps.

The Gaussian sensor-based RHI-G (and RHI-GR) has almost the similar search performance to binary sensor-based RHI-B (and RHI-BR) in terms of mean search time and success rate, but it requires much longer computation time. Thus, RHI-B could be a suitable choice particularly for obstacle environments while RHI-BR showing the reasonable performance could be used instead when the mobile agent has not enough computing power. The RHI-G could be a solution when the computation time does not significantly affect the search performance by using a powerful computer onboard the unmanned ground or sea surface vehicles for not time-critical release/monitoring scenarios.

C. Illustrative run

This section describes an illustrative run of the proposed RHI-BR algorithm which has a reasonably low computational complexity and good search performance. The gas dispersion map is generated by the GRAL model as shown in Fig. 7. The following parameters are used for simulations:

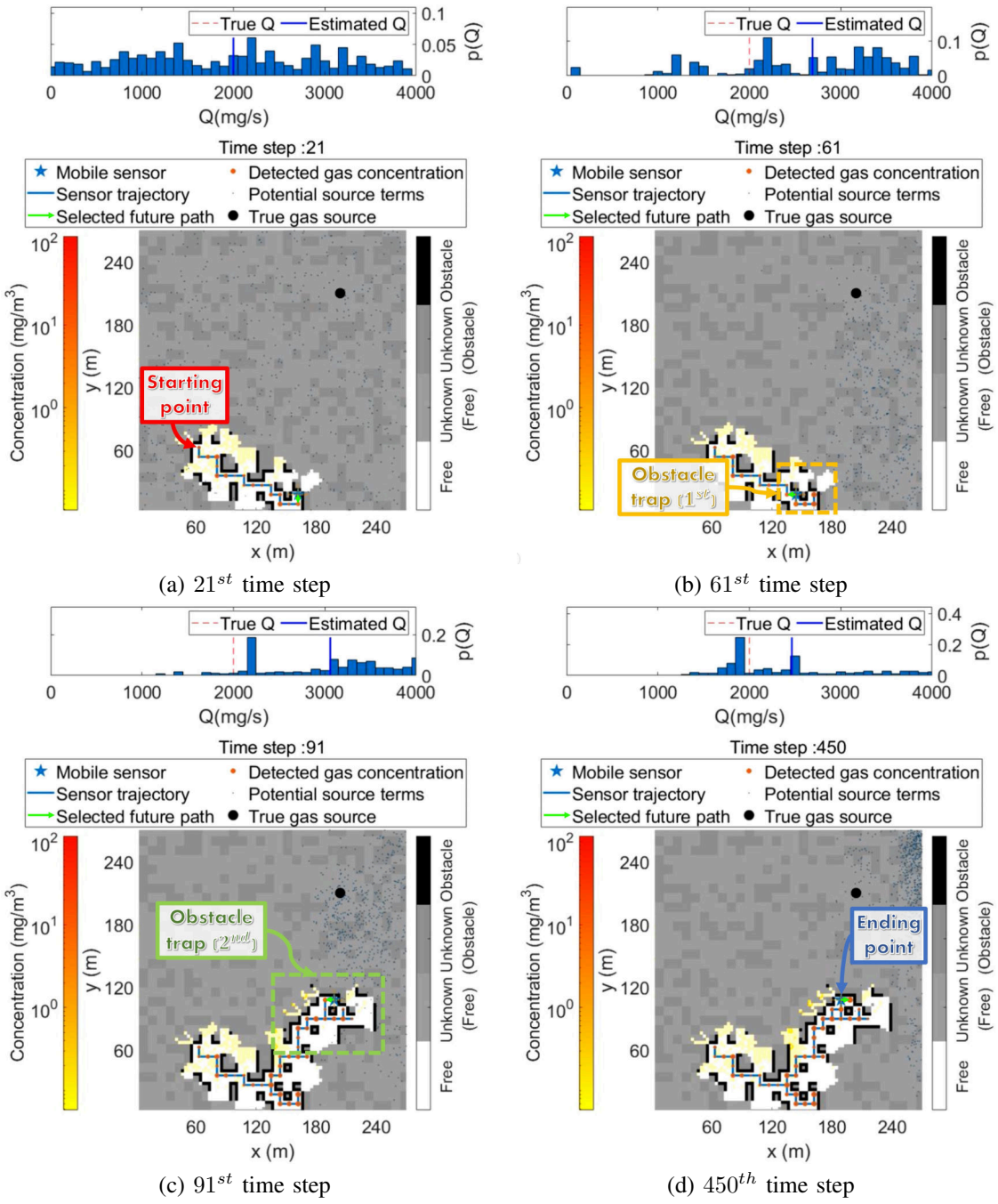
- Search area $A = 270m \times 270m$, wind direction $\phi = 220°$ with velocity $V = 2m/s$, the effective diffusivity parameter $\zeta = 10n^2/s$, release strength of the substance $Q = 7.2kg/h = 2000mg/s$, particle lifetime $\tau = 1000s$, and source location $\mathbf{p}_s = [235m, 235m]^\mathsf{T}$; and
- Feasible control actions, $U = [\uparrow, \downarrow, \leftarrow, \rightarrow]$, and the moving distance of each step $9m$, the standard deviation of environmental and sensor noise are $\sigma_{env} = 1mg/m^3$ and $\sigma_{sen} = 0.2 \cdot \mu(p;\theta)mg/m^3$ respectively, the number of particles $N_p = 1,000$ for the particle filter, the belief ratio $r = 0.5$, binary measurement threshold update parameter $\lambda = 0.8$, and the number of sampled future measurement sequences $N_m = 2$.

The search paths of the greedy Infotaxis and RHI-BR for this scenario are shown in Fig. 11 and Fig. 12, respectively. The probability of the release strength Q is presented in the upper graph as a histogram, where the red dashed line and blue line represent the ground truth and estimated mean of the gas release strength, respectively. The search trajectory and particles are shown in the bottom graph. The blue line is the search path history, and the blue dots are particles from the particle filter. The red dot marks represent the measurement and the bigger size means a higher concentration. The gray environment represents an unknown (unvisited) area while the visited area is covered with white (empty space) and black (walls or obstacles) colors. Note that, the mobile agent is assumed to have a laser scanner (e.g., LiDAR) which provides the range to obstacles within 24 meters radius omnidirectionally. The greedy Infotaxis method presented in Fig. 11 is trapped around obstacles multiple times. Since the estimated source location is behind the obstacles, the mobile agent tries to pass through the building to get better information but the agent is blocked by the building corner; thus, the agent wanders around the building corner for a while and manages to escape sometimes. This behavior occurs quite often in this complex urban environment.

The proposed RHI-BR with the five horizon steps yielded significantly improved performance. RHI-BR can escape quickly from the obstacle traps that the areas marked by green and yellow dashed rectangles. The proposed method enables to obtain better information of the source while avoiding obstacles considering multiple future steps. Thus, the mobile agent can find the source efficiently and estimate the source term correctly as shown in Fig. 12.

D. Performance comparison on realistic conditions

In this section, we conduct the search performance comparison of various algorithms including the proposed approaches using the realistic dispersion map generated by the GRAL model [48]. The RHI-BR approach is compared with a simpler version of the receding horizon algorithm that uses the same feasible paths as the RHI-BR (i.e., not stuck in the environment) but with only a utility value of the first location in the path rather than the summation of predictive utilities along the path. This is termed Infotaxis with RH-path planning. In addition, we compare RHI-BR with three and five horizon steps against state-of-the-art algorithm called IWFA-2 (Infotatxis with forbidden area-2) [30]. We present the outstanding performance of the proposed method, RHI-BR with three and five horizon steps with three sampled measurement sequences, in the obstacle-rich urban environment ($\rho_0 = 0.3$). Each algorithm is tested in ten different maps and a hundred Monte Carlo simulations for each map. The starting location for the



(a) 21^{st} time step

(b) 61^{st} time step

(c) 91^{st} time step

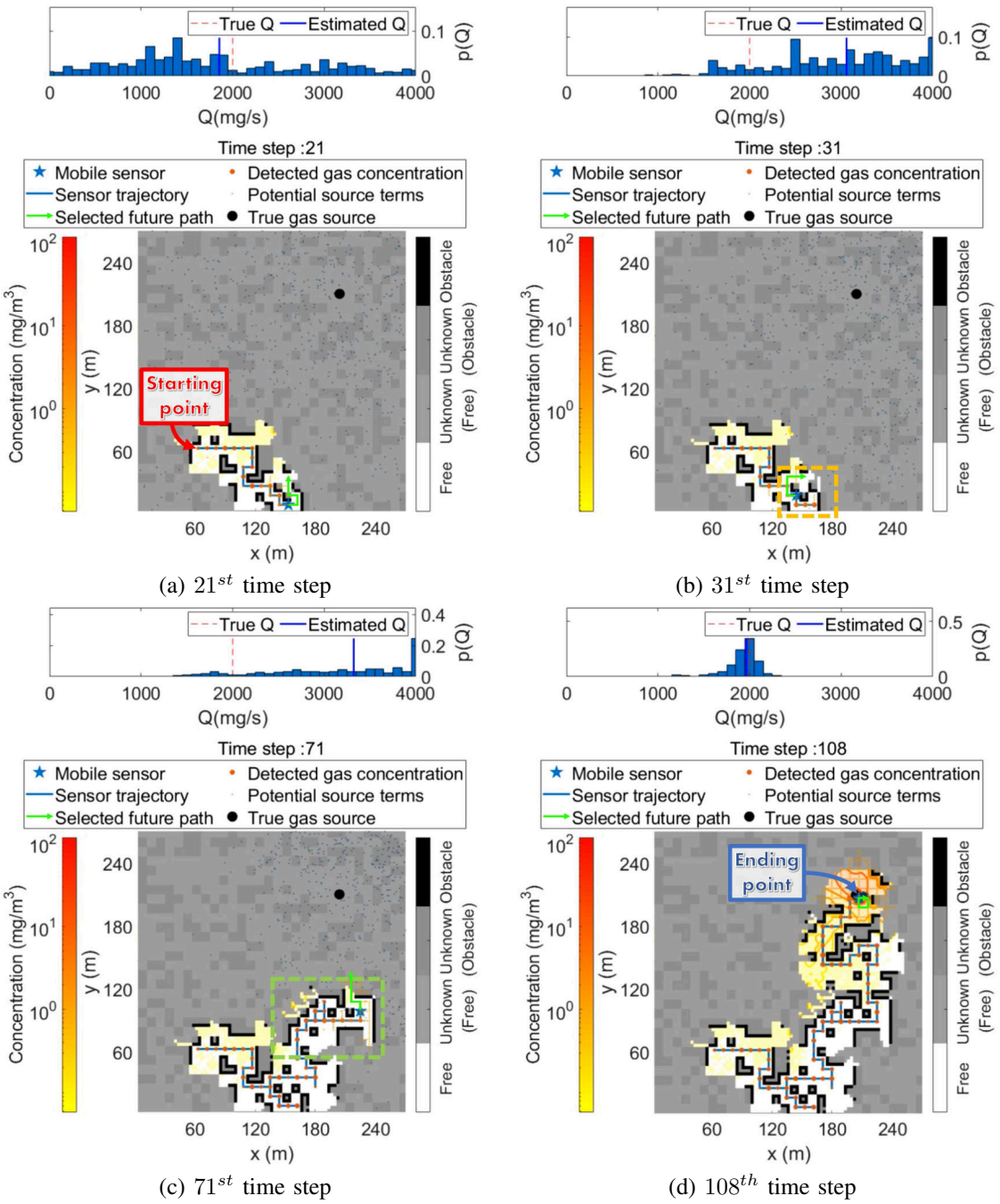
(d) 450^{th} time step

Fig. 11. The result of the greedy Infotaxis source term estimation in a complex obstructed area.

mobile agent is chosen randomly, and for the same map, the random initial locations are saved and used for different algorithms.

The Infotaxis with path planning considers the future path in the same way as the proposed algorithm, but this algorithm uses only the utility that can be obtained at the first location of each path when determining the optimal path. The IWFA-2 avoids revisiting some areas by marking as them forbidden areas while conducting the greedy Infotaxis strategy [30]. Although IWFA-2 outperforms the greedy Infotaxis, this method only considers the geometrical information and it is still based on the greedy approach. Monte Carlo simulation results for the averaged search performance over various maps for each algorithm are shown in Table I.

Note that SR represents the success rate, MST means the total mean search time step, and CT is the computation time. The search performance of Infotaxis with RH-path planning is similar to that of IWFA-2; this is better than Infotaxis but worse than using the sum of predictive utilities in the path (i.e., RHI-BR). Besides, the search performance using the Gaussian sensor model is slightly higher than using the binary sensor model in all cases, but the Gaussian sensor model requires a longer computation time. Note that considering future paths that do not fall into local minima (e.g., dead-



(a) 21^{st} time step (b) 31^{st} time step

(c) 71^{st} time step (d) 108^{th} time step

Fig. 12. The result of the receding horizon Infotaxis source term estimation using the five horizon steps and the three sampled binary sensor measurement sequences in a complex obstructed area.

end) improves the search performance to some extent, but considering the future predictive utilities provide even better performance without a significant increase in computation time. The search performance of RHI-BR (3 steps) is better than IWFA-2 and Infotaxis with Gaussian sensor and five steps horizon length path with less computation time. Although the computing time of RHI-BR (5 steps) is the longest (i.e., 0.05 seconds) among the comparisons, it is fast enough comparing the gas sensing time and moving time. Note that, typically, the gas sensing time and moving time between sensing positions of the mobile agent is the order of a few seconds. Since the true gas dispersion is generated by a realistic GRAL model in these simulations, which is different from the gas dispersion model used by the agent for estimation (Eq. (1)), the success rate of algorithms is worse than that shown in Fig. 9. Although model inconsistency between the ideal gas distribution and GRAL model reduces the search performance to some extent, the success rate of RHI-BR is still above 80% thanks to the robustness of the particle filter and receding-horizon approach. The simulation results of the algorithms that has the best performance among the same algorithms (i.e., Infotaxis, IWFA-2, Infotaxis with Gaussian sensor and five

TABLE I
AVERAGED PERFORMANCE COMPARISON FOR DIFFERENT ALGORITHMS.

	Horizon length	Sensor model	SR [%]	MST [steps]	CT [sec]
Infotaxis	1	Binary	49.4	80.6	0.0011
		Gaussian	51.1	80.4	0.0084
IWFA-2	1	Gaussian	68.0	78.2	0.0084
Infotaxis with RH-path planning	3	Binary	67.5	76.6	0.0018
		Gaussian	68.0	77.5	0.0081
	5	Binary	70.0	78.0	0.0027
		Gaussian	70.5	77.3	0.0091
RHI-BR	3	Binary	81.4	77.0	0.0053
	5	Binary	83.5	74.6	0.0490

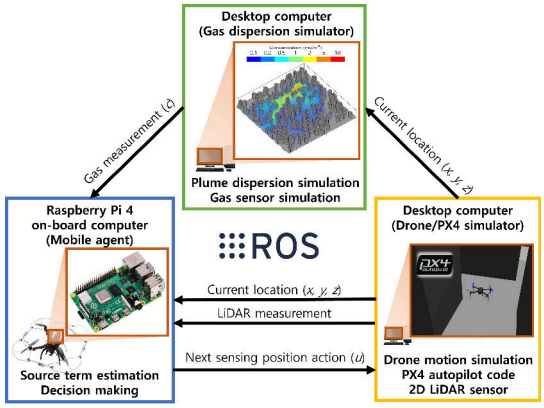


Fig. 14. The hardware-in-the-loop simulation system.

step horizon length path planning, and RHI-BR with 5 steps) for each map are shown in Fig. 13

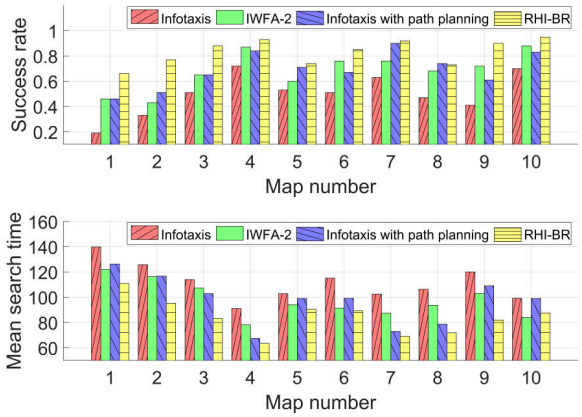


Fig. 13. The performance comparison for different algorithms in various randomly generated maps and GRAL dispersion model.

In 8 map, the proposed algorithm (i.e., RHI-BR) has a slightly lower success rate than Infotaxis with path planning, but it can find the source faster. In 10 map, it moves slightly more than IWFA-2, but it can be seen that it has higher success rate. Except two maps, the proposed RHI-BR algorithm shows the best performance in terms of success rate and number of search moves in all maps.

E. Hardware-in-the-loop simulation in a 3-D environment

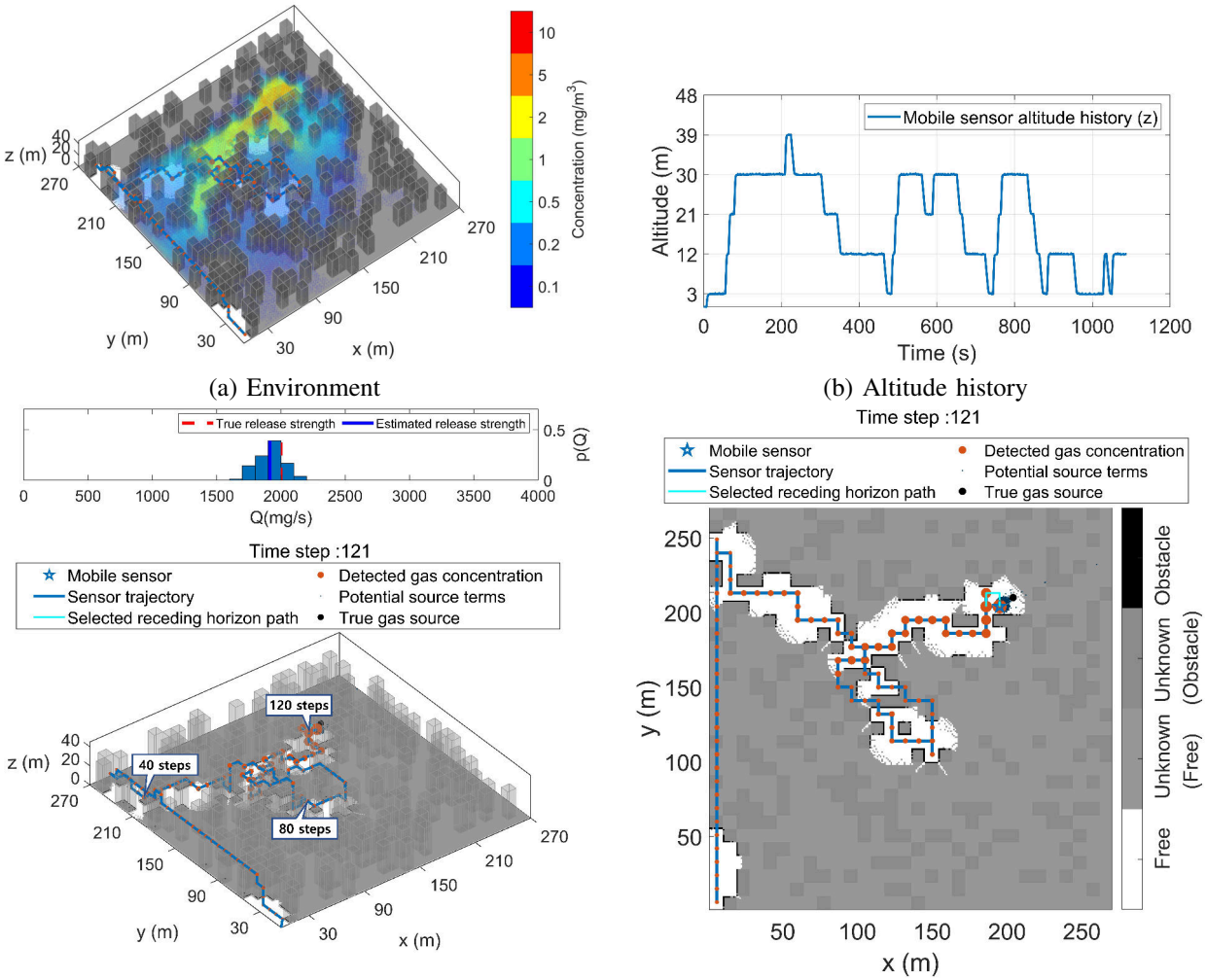
The previous simulations consider only 2-D environments assuming the fixed altitude of the mobile sensor (i.e., UAV). However, if the UAV altitude is far from the height of the source origin, the search performance could be significantly degraded. Thus, in this section, 3-D environments are considered. Besides, to verify the feasibility of the proposed algorithm in real 3-D environments, we perform the hardware-in-the-loop simulation (HILS) (or processor-in-the-loop simulation (PILS) depending on the research community) by using the RaspberryPi 4 board as the onboard computer suitable for a small-sized multirotor drone. Note that, although there are some real outdoor drone experiments in an open space environment (i.e., without obstacles) [7], [9], [10] including our previous work [8], flight experiments in obstacle-rich environments are not yet reported due to the difficulty of making actual gas dispersion strong enough to be detected by a real gas sensor in such complex environments. So, here, we use the HILS system instead of real flight as illustrated in Fig. 14.

In the HILS system, the onboard computer performs source term estimation using the particle filter and decision making using RHI-BR. There are two environmental simulators: one for simulating realistic gas dispersion in an obstacle-rich environment using the GRAL simulator (explained previously), and the other for simulating the realistic dynamics and control of a drone with the PX4 autopilot code in a Gazebo simulator. The drone model that is used in Gazebo is IRIS quadrotor UAV. The current state is calculated through an extended Kalman filter (EKF) using IMU and GNSS sensors in the simulation. The autopilot calculates motor power based on its current state for stabilization and control. The Gazebo simulation calculates the wind lift forces generated by rotating motors and propellers. The PID controller is adopted to control the UAV velocity.

The gas dispersion simulator provides noisy gas concentration measurements to the drone onboard computer, and the drone simulator gives the current position and LiDAR measurements. The HILS system is run on the robot operating system (ROS).

For 3-D HILS, the previous gas dispersion model in Eq. (1) is extended to 3-D by replacing the 2-D location of the source and the UAV with 3-D ones in the equation. The feasible action set is changed to $\mathbf{u} = [\uparrow, \downarrow, \rightarrow, \leftarrow, \Uparrow, \Downarrow]$ where $\Uparrow$ and $\Downarrow$ represent the altitude change and the movement distance between time steps is fixed at 9m for all six directions. The height of the true source is 12m and the initial height of the UAV is set to 3m. The wind is assumed to be blown only horizontally. The building generation probability $\rho_0 = 0.3$ and the other environmental parameters are the same as the previous 2-D simulation.

The HILS result for an illustrative run is shown in Fig. 15. The proposed algorithm successfully estimates the source term in a 3-D environment by including altitude control in the feasible action set. The total search time for 3-D HILS can



(a) Environment

(b) Altitude history

(c) Search path (3-D view)

(d) Search path (top view)

Fig. 15. The illustrative run of hardware-in-the-loop simulation.

TABLE II
THE COMPUTATION TIME FOR DIFFERENT ENVIRONMENTS.

	Desktop (MATLAB, 2D)	Desktop (MATLAB, 3D)	Onboard RPI (C++, 2D)	Onboard RPI (C++, 3D)
Computation Time	0.0058 sec	0.0244 sec	0.0994 sec	0.493 sec

be found in Fig. 15 (b). The computation times for RHI-BR (3 steps) in different environment setup are shown in Table II. As the gas sensing generally takes a couple of seconds while hovering over the sensing location per search time step [7], [8] as explained in Section II, the computation time of about half second for the 3-D onboard case is still not that long for real-time applications comparing sensing time (about three seconds) and movement time (about five seconds). Besides, a more powerful onboard computer could be used if needed.

VI. Conclusions and Future Work

Information-theoretic source search planning was formulated as receding horizon path planning to maximize the entropy reduction for the estimated source term. The binary sensor model and random sampling method were introduced to reduce the computational complexity of the receding horizon approach for real-time decision making. The proposed search method shows much better performance in an obstacle-rich environment compared with the existing greedy Infotaxis and a state-of-the-art IWFA-2 algorithm for source search in an obstacle-rich environment.

There are several future extensions to improve the performance of the proposed approach and make it more practical. First, considering more feasible actions for each step (e.g., diagonal motions, different step sizes between sensing points) or trajectory planning with a variable agent speed could increase the search performance. However, to utilize a larger number of actions or optimization, the decision making

(i.e., path planning) algorithm should be further improved so that the computational time does not increase exponentially. Besides, if the movement distance for each step is variable, it can be included in the utility function to minimize the total travel cost along with the entropy reduction. An efficient waypoint-sampling method in a continuous domain, such as the rapidly-exploring random tree (RRT) [49], [50], could be combined with the proposed receding horizon-based Infotaxis strategy to improve the search performance. Second, employing a more complex and realistic dispersion model (e.g., considering a time-varying release rate) could improve the estimation performance rather than analytical gas dispersion models. However, if we employ a complex model such as the CFD model or puff-based model, the computational burden could be greatly increased when predicting the expected gas concentration particularly for the particle filter; thus, a trade-off between the fidelity of the model and computational complexity needs to be carefully considered. Third, in this study, the monotonically increasing binary sensor threshold and fixed decay factor is introduced, but these terms should be further investigated in case of the changing source parameters or strong turbulent and unstable environment conditions. We also plan to develop the source search algorithm using the deep reinforcement learning approach to calculate and predict complex gas dispersion and decide optimal search paths in real-time. Lastly, a field experiment using the proposed method in obstacle-rich environments will be considered for future work.

ACKNOWLEDGEMENTS

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education (2020R1A6A1A03040570) and National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (2020R1F1A1049066).

REFERENCES

[1] M. Hutchinson, H. Oh, and W.-H. Chen, "A review of source term estimation methods for atmospheric dispersion events using static or mobile sensors," *Information Fusion*, vol. 36, pp. 130–148, 2017.

[2] S. K. Singh and R. Rani, "A least-squares inversion technique for identification of a point release: Application to fusion field trials 2007," *Atmospheric Environment*, vol. 92, pp. 104–117, 2014.

[3] P. Sujit and D. Ghose, "Search using multiple UAVs with flight time constraints," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 40, no. 2, pp. 491–509, 2004.

[4] D. Erdos, A. Erdos, and S. E. Watkins, "An experimental UAV system for search and rescue challenge," *IEEE Aerospace and Electronic Systems Magazine*, vol. 28, no. 5, pp. 32–37, 2013.

[5] S. M. Esmailifar and F. Saghafi, "Moving target localization by cooperation of multiple flying vehicles," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 51, no. 1, pp. 739–746, 2015.

[6] P. Boström-Rost, D. Axehill, and G. Hendeby, "Sensor management for search and track using the poisson multi-bernoulli mixture filter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 57, no. 5, pp. 2771–2783, 2021.

[7] M. Hutchinson, C. Liu, P. Thomas, and W.-H. Chen, "Unmanned aerial vehicle-based hazardous materials response: Information-theoretic hazardous source search and reconstruction," *IEEE Robotics & Automation Magazine*, vol. 27, no. 3, pp. 108–119, 2020.

[8] M. Park, S. An, J. Seo, and H. Oh, "Autonomous source search for UAVs using gaussian mixture model-based Infotaxis: Algorithm and flight experiments," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 57, no. 6, pp. 4238–254, 2021.

[9] J. Valente, S. Munniks, I. de Man, and L. Kooistra, "Validation of a small flying E-nose system for air pollutants control: A plume detection case study from an agricultural machine," in *2018 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, 2018, pp. 1993–1998.

[10] X. He, J. R. Bourne, J. A. Steiner, C. Mortensen, K. C. Hoffman, C. J. Dudley, B. Rogers, D. M. Cropek, and K. K. Leang, "Autonomous chemical-sensing aerial robot for urban/suburban environmental monitoring," *IEEE Systems Journal*, vol. 13, no. 3, pp. 3524–3535, 2019.

[11] D. Shin, Y. Song, J. Oh, and H. Oh, "Nonlinear disturbance observer-based standoff target tracking for small fixed-wing UAVs," *International Journal of Aeronautical and Space Sciences*, vol. 22, no. 1, pp. 108–119, 2021.

[12] A. Celani and M. Vergassola, "Bacterial strategies for Chemotaxis response," *Proceedings of the National Academy of Sciences*, vol. 107, no. 4, pp. 1391–1396, 2010.

[13] N. Voges, A. Chaffiol, P. Lucas, and D. Martinez, "Reactive searching and Infotaxis in odor source localization," *PLOS Computational Biology*, vol. 10, no. 10, p. e1003861, 2014.

[14] P. P. Neumann, V. Hernandez Bennetts, A. J. Lilienthal, M. Bartholmai, and J. H. Schiller, "Gas source localization with a micro-drone using bio-inspired and particle filter-based algorithms," *Advanced Robotics*, vol. 27, no. 9, pp. 725–738, 2013.

[15] D. J. Harvey, T.-F. Lu, and M. A. Keller, "Comparing insect-inspired chemical plume tracking algorithms using a mobile robot," *IEEE Transactions on Robotics*, vol. 24, no. 2, pp. 307–317, 2008.

[16] J. R. Bourne, E. R. Pardyjak, and K. K. Leang, "Coordinated bayesian-based bioinspired plume source term estimation and source seeking for mobile robots," *IEEE Transactions on Robotics*, vol. 35, no. 4, pp. 967–986, 2019.

[17] M. Vergassola, E. Villermaux, and B. I. Shraiman, "Infotaxis as a strategy for searching without gradients," *Nature*, vol. 445, no. 7126, p. 406, 2007.

[18] B. Ristic, A. Skvortsov, and A. Gunatilaka, "A study of cognitive strategies for an autonomous search," *Information Fusion*, vol. 28, pp. 1–9, 2016.

[19] M. Hutchinson, H. Oh, and W.-H. Chen, "Entrotaxis as a strategy for autonomous search and source reconstruction in turbulent conditions," *Information Fusion*, vol. 42, pp. 179–189, 2018.

[20] M. Park and H. Oh, "Cooperative information-driven source search and estimation for multiple agents," *Information Fusion*, vol. 54, pp. 72–84, 2020.

[21] N. E. Humphries, N. Queiroz, J. R. Dyer, N. G. Pade, M. K. Musyl, K. M. Schaefer, D. W. Fuller, J. M. Brunnschweiler, T. K. Doyle, J. D. Houghton *et al.*, "Environmental context explains Lévy and Brownian movement patterns of marine predators," *Nature*, vol. 465, no. 7301, p. 1066, 2010.

[22] L. Champagne, R. G. Carl, and R. Hill, "Agent models ii: search theory, agent-based simulation, and U-boats in the bay of biscay," in *Proceedings of the 35th Conference on Winter Simulation: Driving Innovation*. Winter Simulation Conference, 2003, pp. 991–998.

[23] Q. Lu and Q.-L. Han, "Cooperative control of a multi-robot system for odor source localization," in *IECON 2011 - 37th Annual Conference of the IEEE Industrial Electronics Society*, 2011, pp. 2359–2364.

[24] M. Park and H. Oh, "Information-driven autonomous search and source reconstruction using cooperative mobile sensors," in *AIAA Scitech 2019 Forum*, 2019, p. 0397.

[25] R. Khodayi-mehr, W. Aquino, and M. M. Zavlanos, "Model-based active source identification in complex environments," *IEEE Transactions on Robotics*, vol. 35, no. 3, pp. 633–652, 2019.

[26] Y. Zhao, B. Chen, Z. Zhu, F. Chen, Y. Wang, and D. Ma, "Entrotaxis-jump as a hybrid search algorithm for seeking an unknown emission source in a large-scale area with road network constraint," *Expert Systems with Applications*, p. 113484, 2020.

[27] J.-G. Li, J. Yang, J. Liu, G.-D. Lu, and L. Yang, "Odor-source searching using a mobile robot in time-variant airflow environments with obstacles," in *Proceedings of the 33rd Chinese Control Conference*. IEEE, 2014, pp. 8559–8564.

[28] R. Zou, M. Zhang, V. Kalivarapu, E. Winer, and S. Bhattacharya, "Particle swarm optimization for source localization in environment with obstacles," in *2014 IEEE International Symposium on Intelligent Control (ISIC)*, 2014, pp. 1602–1607.

[29] Z. Liu and T.-F. Lu, "Multiple robots plume-tracing in open space obstructed environments," in *2009 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, 2009, pp. 2433–2439.

[30] Y. Zhao, B. Chen, Z. Zhu, F. Chen, Y. Wang, and Y. Ji, "Searching the diffusive source in an unknown obstructed environment by cognitive strategies with forbidden areas," *Building and Environment*, p. 107349, 2020.

[31] M. Hutchinson, C. Liu, and W.-H. Chen, "Source term estimation of a hazardous airborne release using an unmanned aerial vehicle," *Journal of Field Robotics*, vol. 36, no. 4, pp. 797–817, 2019.

[32] B. Ristic, A. Gunatilaka, and R. Gailis, "Localisation of a source of hazardous substance dispersion using binary measurements," *Atmospheric Environment*, vol. 142, pp. 114–119, 2016.

[33] W. Jung and H. Bang, "Fault and failure tolerant model predictive control of quadrotor UAV," *International Journal of Aeronautical and Space Sciences*, pp. 1–13, 2021.

[34] E. T. Jaynes, *Probability Theory: The Logic of Science*. Cambridge, U.K.: Cambridge University Press, 2003.

[35] E. Yee, "Automated computational inference engine for bayesian source reconstruction: Application to some detections/non-detections made in the CTBT international monitoring system," *Applied Mathematical Sciences*, vol. 11, no. 32, pp. 1581–1618, 2017.

[36] G. C. Efthimiou, I. V. Kovalets, A. Venetsanos, S. Andronopoulos, C. D. Argyropoulos, and K. Kakosimos, "An optimized inverse modelling method for determining the location and strength of a point source releasing airborne material in urban environment," *Atmospheric Environment*, vol. 170, pp. 118–129, 2017.

[37] S. Chakraborty, "Generating discrete analogues of continuous probability distributions-a survey of methods and constructions," *Journal of Statistical Distributions and Applications*, vol. 2, no. 1, p. 6, Aug 2015.

[38] J.-G. Li, Q.-H. Meng, Y. Wang, and M. Zeng, "Odor source localization using a mobile robot in outdoor airflow environments with a particle filter algorithm," *Autonomous Robots*, vol. 30, no. 3, pp. 281–292, 2011.

[39] H. Hajieghrary, M. A. Hsieh, and I. B. Schwartz, "Multi-agent search for source localization in a turbulent medium," *Physics Letters A*, vol. 380, no. 20, pp. 1698–1705, 2016.

[40] J. A. Gubner, *Probability and Random Processes for Electrical and Computer Engineers*. Cambridge, U.K.: Cambridge University Press, 2006.

[41] B. Ristic, S. Arulampalam, and N. Gordon, *Beyond the Kalman Filter: Particle Filters for Tracking Applications*. Norwood, MA: Artech house, 2004.

[42] G. E. Collins, J. R. Riehl, and P. S. Vegdahl, "A UAV routing and sensor control optimization algorithm for target search," in *Unmanned Systems Technology IX*, vol. 6561. International Society for Optics and Photonics, 2007, p. 65610D.

[43] C. Liu and J. K. Hedrick, "Model predictive control-based target search and tracking using autonomous mobile robot with limited sensing domain," in *2017 American Control Conference (ACC)*, 2017, pp. 2937–2942.

[44] A. Ryan and J. K. Hedrick, "Particle filter based information-theoretic active sensing," *Robotics and Autonomous Systems*, vol. 58, no. 5, pp. 574–584, 2010.

[45] G. Grisetti, C. Stachniss, and W. Burgard, "Improved techniques for grid mapping with Rao-Blackwellized particle filters," *IEEE Transactions on Robotics*, vol. 23, no. 1, pp. 34–46, 2007.

[46] D. Ben-Avraham and S. Havlin, *Diffusion and Reactions in Fractals and Disordered Systems*. Cambridge, U.K.: Cambridge University Press, 2000.

[47] S. R. Finch, *Mathematical Constants*. Cambridge, U.K.: Cambridge University Press, 2003.

[48] D. Oettl, "Evaluation of the revised lagrangian particle model GRAL against wind-tunnel and field observations in the presence of obstacles," *Boundary-Layer Meteorology*, vol. 155, no. 2, pp. 271–287, 2015.

[49] K. Yang, "Anytime synchronized-biased-greedy rapidly-exploring random tree path planning in two dimensional complex environments," *International Journal of Control, Automation and Systems*, vol. 9, no. 4, p. 750, 2011.

[50] G. Noh, J. Park, D. Han, and D. Lee, "Selective goal aiming rapidly exploring random tree path planning for UAVs," *International Journal of Aeronautical and Space Sciences*, vol. 22, no. 6, pp. 1397–1412, 2021.

Minkyu Park received the B.Sc. degree in mechanical system design and manufacturing, in 2016, from the Ulsan National Institute of Science and Technology (UNIST), South Korea. He is currently working toward the Ph.D degree at the Autonomous Systems Laboratory, UNIST. His research interests include autonomous search, sensor management, Bayesian estimation, information theory and mobile robot application for environmental monitoring.

Pawel Ladosz received the Meng (hons) degree in aerospace engineering in 2014 from Manchester University and a PhD degree in 2019 from Loughborough University. He is currently Research Assistant Professor at Ulsan National Institute of Science and Technology (UNIST). His research interests include reinforcement learning, robotics, optimization and deep learning.

Jongyun Kim received the B.Sc. degree in mechanical system design and manufacturing in 2017 and M.Sc. degree in mechanical engineering in 2020 from Ulsan National Institute of Science and Technology (UNIST), South Korea. He is currently pursuing his Ph.D. degree at the School of Aerospace, Transport and Manufacturing, Cranfield University, United Kingdom. His research interests include multi-robot systems, real-time decision making, communication relay, reinforcement learning, and stochastic optimization.

Hyondong Oh (Senior Member, IEEE) received BSc and MSc degrees in aerospace engineering from Korea Advanced Institute of Science and Technology (KAIST), South Korea, in 2004 and 2010, respectively, and a PhD in autonomous surveillance and target tracking guidance of multiple UAVs from Cranfield University, U.K., in 2013. He was a lecturer in the field of autonomous unmanned vehicles at Loughborough University, U.K., from 2014 to 2016. He is currently an associate professor with the Department of Mechanical Engineering, Ulsan National Institute of Science and Technology (UNIST). His research interests include autonomy and decision making, cooperative control, path planning, nonlinear guidance and control, and estimation and sensor/information fusion for unmanned systems.

2022-06-21

Receding horizon-based infotaxis with random sampling for source search and estimation in complex environments

Park, Minkyu

IEEE

Park M, Ladosz P, Kim J, Oh H. (2023) Receding horizon-based infotaxis with random sampling for source search and estimation in complex environments. IEEE Transactions on Aerospace and Electronic Systems, Volume 59, Issue 1, February 2023, pp. 591-609

https://doi.org/10.1109/TAES.2022.3184923