# Preprints.org

# Robotic Odor Source Localization via Vision and Olfaction Fusion Navigation Algorithm

Sunzid Hassan , Lingxiao Wang * , Khan Raqib Mahmud

*Article*

# Robotic Odor Source Localization via Vision and Olfaction Fusion Navigation Algorithm

**Sunzid Hassan [1], Lingxiao Wang [2],\*  and Khan Raqib Mahmud [1]**

[1]  Department of Computer Science, Louisiana Tech University, 201 Mayfield Ave, Ruston, LA 71272, US; sha040@latech.edu (S.H.); krm070@email.latech.edu (K.M.)

[2]  Department of Electrical Engineering, Louisiana Tech University, 201 Mayfield Ave, Ruston, LA 71272, US; lwang@latech.edu

\*  Correspondence: lwang@latech.edu; Tel.: +1-318-257-2758

**Abstract:** Robotic odor source localization (OSL) is a technology that enables mobile robots or autonomous vehicles to find an odor source in unknown environments. An effective navigation algorithm that guides the robot to approach the odor source is the key to successfully locating the odor source. While traditional OSL approaches primarily utilize an Olfaction-only strategy, guiding robots to find the odor source by tracing emitted odor plumes, our work introduces a fusion navigation algorithm that combines both vision and olfaction-based techniques. This hybrid approach addresses challenges such as turbulent airflow, which disrupts olfaction sensing, and physical obstacles inside the search area, which may impede vision detection. In this work, we propose a hierarchical control mechanism that dynamically shifts the robot's search behavior among four strategies: crosswind maneuver, Obstacle-avoid Navigation, Vision-based Navigation, and Olfaction-based Navigation. Our methodology includes a custom-trained deep-learning model for visual target detection and a moth-inspired algorithm for Olfaction-based navigation. To assess the effectiveness of our approach, we implemented the proposed algorithm on a mobile robot in a search environment with obstacles. Experimental results demonstrate that our Vision and Olfaction Fusion algorithm significantly outperforms Vision-only and Olfaction-only methods, reducing average search time by 54% and 30%, respectively.

**Keywords:** odor source localization; moth-inspired algorithm; computer Vision-based Navigation; robot operating system; multi-modal robotics.

---

## 1. Introduction

Sensory systems like olfaction, vision, audition, etc., allow animals to interact with the external environment. Among these, olfaction is the oldest sensory system to evolve in organisms [1]. Olfaction allows organisms with receptors for the odorant to identify food, potential mating partners, dangers, and enemies [2]. In some nocturnal mammals like mice, as much as five percent of the genome is devoted to olfaction [3]. Similar to animals, a mobile robot integrated with a chemical sensor can detect odors in the external environment. Robotic Odor source localization (OSL) is the technology that allows robots to utilize olfaction sensory inputs to navigate toward an unknown target odor source in the given environment [4]. It has important applications including monitoring wildfires [5], locating air pollution [6], locating chemical gas leaks [7], locating unexploded mines and bombs [8], locating underground gas leaks [9], and marine surveys such as finding hydrothermal vents [10], etc.

Locating an unknown odor source requires an effective OSL algorithm guiding the robot based on sensor observations. Current OSL algorithms include bio-inspired methods that imitate animal olfactory behaviors, engineering-based methods that rely on mathematical models to estimate potential odor source locations and machine learning-based methods that use a trained model to guide the robot toward the odor source. The typical bio-inspired method includes moth-inspired algorithm that imitates male months mate-seeking behaviors [11], where a robotic agent will follow a 'surge/casting' model [12] to reach the odor source. Typical engineering-based methods includes the Particle Filter

algorithm [13], where the robot will use historic olfaction reading to predict the odor source location. Finally, typical machine learning-based OSL methods include deep supervised and reinforcement learning-based methods.

All of these approaches rely on olfaction (e.g., chemical and airflow) sensing to detect and navigate to the given odor source. However, approaches that rely solely on olfaction sensing struggle in turbulent airflow environments. In contrast, animals that operate in complex airflow environments often rely on multiple sensory systems like olfaction and vision for odor source localization. For example, humans often recognize the presence of an odor source of interest with olfaction (e.g., smelling a barbecue), and locate and navigate to the odor source using vision (e.g., locating the barbecue shop with vision). If there is no valid vision of the odor source, we may search for the source using olfaction sensing (e.g., moving towards the direction of greater odor concentration or against the direction of wind flow). Similarly, a robot with both olfaction and vision sensing capabilities (e.g., with a camera and chemical sensor) can find an unknown odor source more efficiently, compared to olfaction-only OSL navigation methods. Thus, this project departs from the existing OSL navigation methods in utilizing both robotic vision and olfaction for searching the odor source location. The core of this project involves designing an algorithm that utilizes both vision and olfaction sensing for locating an unknown odor source location.

The project proposes an effective sensor fusion approach that utilizes a vision method and bio-mimicking olfaction method to guide the robot toward an unknown odor source in a real-world obstacle-ridden search area with both laminar and turbulent airflow setups. Figure 1 shows the proposed method, where we show the developed robot platform equipped with vision and olfaction sensors. The vision sensors include a camera, and the olfaction sensors include a chemical detector and anemometer. It also includes a Laser Distance Sensor (LDS) for obstacle detection. The sensor observations are transmitted to a decision-making model, which is implemented in a remote computer. The model selects Obstacle-avoid Navigation, Vision-based Navigation, or Olfaction-based Navigation behavior based on the sensor readings. In the proposed decision-making model, the robotic vision is achieved by a deep-learning vision model, and the robotic olfaction model is based on a bio-mimicking moth-inspired algorithm. Based on the current sensor reading, the active search behavior will calculate the robot heading commands, guiding the robot to approach the odor source location. Finally, the robot executes the heading command, collects new sensor readings at the new location, and repeats the loop until the odor source is detected.
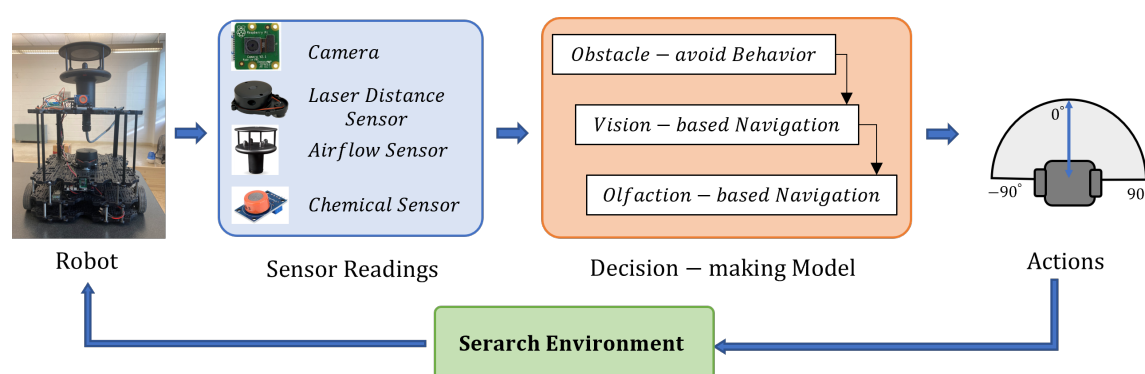


**Figure 1.** Flow diagram of the proposed method for OSL experiment. We utilized the Turtlebot3 robot platform. We equipped it with a camera, Laser Distance Sensor, Airflow sensor, Chemical sensor, etc. The robot utilizes 3 navigation behaviors - Obstacle-avoid Navigation, Vision-based Navigation, and Olfaction-based Navigation to output robot heading and linear velocity.

In order to test the performance of our proposed Vision and Olfaction Fusion Navigation algorithm, we conducted 30 real-world OSL experiments using Olfaction-only Navigation algorithm, Vision-only Navigation Algorithm, and the proposed Vision and Olfaction Fusion Navigation

algorithms in both laminar and turbulent airflow environments. Contributions of this work can be summarized as:

1. Introduce vision as an additional sensing modality for odor source localization. For vision sensing, We trained a deep-learning-based computer vision model to detect odor sources from emitted visible plumes.
2. Develop a multimodal Vision and Olfaction Fusion Navigation algorithm with Obstacle-avoid Navigation capabilities for OSL tasks.
3. Compare the search performance of Olfaction-only and Vision-only navigation algorithms with the proposed Vision and Olfaction Fusion Navigation algorithm in a real-world search environment with obstacles and turbulent airflow setups.

In the remaining of this paper, Section 2 reviews the recent progress of olfactory-based navigation algorithms; Section 3 reviews technical details of the proposed OSL algorithm; Section 4 presents details of performing the real-world experiments; Section 5 includes a discussion on the future research direction based on this work; and finally, Section 6 includes overall conclusion of the work.

## 2. Related Works

Research on Robotic Odor Source Localization (OSL) has gained significant attention in recent decades [14]. Technological advancements in robotics and autonomous systems have made it possible to deploy mobile robots for locating odor or chemical sources. Designing algorithms that mimic the navigation method of biological organisms is a typical approach in robotic odor source localization research. Organisms across various sizes rely on scent for locating objects. Whether it's a bacterium navigating an amino acid gradient or a wolf tracking down prey, the ability to follow odors can be crucial for survival.

Chemotaxis is the simplest odor source localization approach in biological organisms, where they rely only on olfaction for navigation. For example, bacteria exhibit chemotaxis by adjusting their movement in response to changes in chemical concentration. When they encounter higher levels of an appealing chemical, their likelihood of making temporary turns decreases, promoting straighter movement. Conversely, in the absence of a gradient or when moving away from higher concentrations, the default turning probability is maintained [15]. This simple algorithm enables single-celled organisms to navigate along a gradient of attractive chemicals through a guided random walk. Nematodes [16] and crustaceans [17] also, follow Chemotaxis-based odor source localization. Early attempts at robotic OSL focused on employing such simple gradient following chemotaxis algorithms. These methods utilized a pair of chemical sensors on plume-tracing robots, directing them to steer towards higher concentration measurements [18]. Several early studies [19–22] validated the effectiveness of chemotaxis in laminar flow environments, characterized by low Reynolds numbers. However, in turbulent flow environments with high Reynolds numbers, alternative methods were proposed, drawing inspiration from both complex biological and engineering principles.

Odor-gated anemotaxis navigation is a more complex odor source localization method that utilizes senses of both odor and airflow for navigation. Moths [23–25], birds [26,27], etc. organisms follow this type of navigation. In particular, mimicking the mate-seeking behavior of male moths led to the development of the moth-inspired method in robotic odor source localization. This method was successfully applied in various robotic OSL scenarios [28] .Additionally, diverse bio-inspired search strategies like zigzag, spiral, fuzzy-inference, and multi-phase exploratory approaches have been introduced [29] in odor-gate anemotaxis-based solutions. Recent bio-inspired OSL navigation methods also aimed to make the search environment more complicated. For instance, [30] proposed a 3-dimensional (3-D) moth-inspired OSL search strategy that utilized cross-wind Lévy Walk, spiraling and upwind surge.

Engineering-based methods take a different approach than bio-mimicking algorithms, relying on mathematical models for estimating odor source locations. These methods are often times known

as Infotaxis [31]. These methods involve constructing source probability maps, dividing the search area into regions, and assigning probabilities indicating the likelihood of containing the odor source. Algorithms for constructing such maps include Bayesian inference, particle filters , stochastic mapping [32], source term estimation [33], information-based search [34], partially observable Markov decision processes [35], etc. Subsequently, robots are guided towards the estimated source via path planning algorithms such as artificial potential fields, A-star [36,37]. These models also rely on olfaction sensing for estimating the odor source.

Deep Learning (DL) based methods are increasingly utilized for OSL experiments. Recent developments involve the use of Deep Neural Networks (DNNs) to predict gas leak locations from stationary sensor networks or employing reinforcement learning for plume tracing strategies. For instance, Kim et al. [38] trained an RNN to predict potential odor source locations using data from stationary sensor networks obtained through simulation. Hu et al. [39] presented a plume tracing algorithm based on model-free reinforcement learning, utilizing the deterministic policy gradient to train an actor-critic network for Autonomous Underwater Vehicle (AUV) navigation. Wang et al. [40] trained an adaptive neuro-fuzzy inference system (ANFIS) to solve the OSL problem in simulations, yet real-world validations are necessary to confirm its efficacy. In summary, despite the promising potential of DL technologies, their application in solving OSL problems is still in its early stages and warrants further research. Most DL-based methods are validated in virtual environments through simulated flow fields and plume distributions, necessitating real-world implementations to validate their effectiveness.

Fusing vision with olfaction for odor source localization task is common in complex organisms like mice [41,42]. Humans also use vision as a primary sensor for odor source navigation tasks. However, very few works have utilized vision for OSL tasks. Recent advances in computer vision techniques can allow robots to use vision as an important sensing capability for detecting visible odor sources or plumes. The added advantage of vision is that it can allow robots to navigate to odor sources without being affected by sparse odor plumes or turbulent airflow in the navigation path. The main contribution of this paper is designing a dynamic Vision and Olfaction Fusion Navigation algorithm for odor source localization in an obstacle-ridden turbulent airflow environment.

## 3. Materials and Methods

### 3.1. Overview of the Proposed OSL Algorithm

Figure 2 shows the flow diagram of the proposed navigation algorithm. In this work, the initial robot search behavior is the 'Crosswind maneuver' behavior, where the robot moves cross-wind to detect initial odor plumes. If the robot encounters obstacles in its surroundings, it switches to the 'Obstacle-avoid Navigation' behavior, where the robot will move around to avoid obstacles. During the robot maneuver, the robot seeks valid visual and olfactory detection. If the robot obtains a valid visual detection, it employs Vision-based Navigation to approach the odor source location. Similarly, if the robot obtains sufficient olfactory detection, it employs Olfaction-based Navigation algorithm. If the robot is in the vicinity of the odor source, it is considered as the source declaration, i.e., the end of the search. Otherwise, the robot returns to the default 'Crosswind maneuver' behavior and repeats the above process.

In the following section, we present the design of the aforementioned search behaviors, including Crosswind maneuver (Subsection 3.2), Obstacle-avoid Navigation (Subsection 3.3), Vision-based Navigation (Subsection 3.4), and Olfactory-based Navigation (Subsection 3.5).
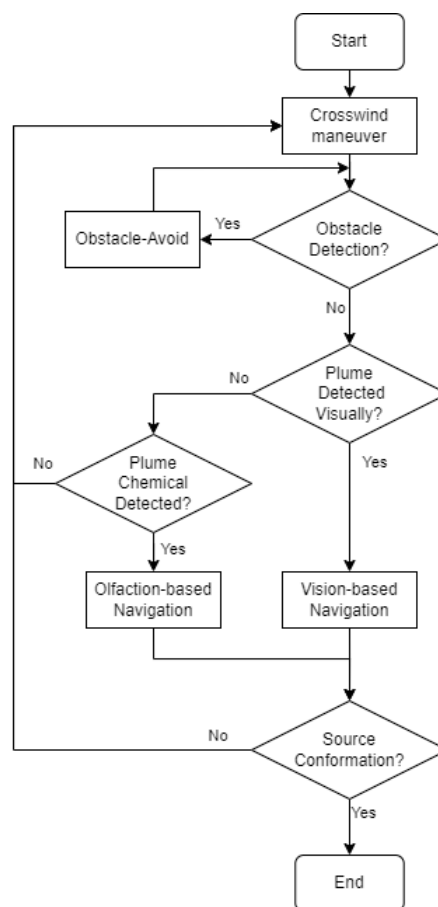
**Figure 2.** The flow diagram of the proposed OSL algorithm. There are four navigation behaviors, including 'Crosswind maneuver', 'Obstacle-avoid Navigation', 'Vision-based Navigation', and 'Olfaction-based Navigation'.

### 3.2. Crosswind maneuver Behavior

In an OSL task, the robot does not have any prior information on the odor source location. Thus, we define a 'Crosswind maneuver' behavior, as the default behavior, directing the robot to find initial odor plume detection or re-detect odor plumes when valid vision and olfaction observations are absent. Crosswind movement, where the robot heading is perpendicular to the wind direction, increases the chance of the robot detecting odor plumes [43]. Denote that the wind direction in the inertial frame is $\phi$, thus, the robot heading command in the 'Crosswind maneuver' behavior can be defined as:

$$\psi_c = \phi_{Inertial} + 90. \tag{1}$$

Besides, it is worth mentioning that we set the robot's linear speed as a constant and only changed the heading commands in the 'Crosswind maneuver' behavior to simplify the robot control problem and save search time.

### 3.3. Obstacle-avoid Navigation Behavior

The 'Obstacle-avoid Navigation' behavior is activated when the robot moves close to an obstacle object within the search environment, which directs the robot to move around and avoid the obstacles. In this work, the robot employs a Laser Distance Sensor (LDS) to measure the distances from the robot to any obstacles in five surrounding angles as presented in Figure 3. Specifically, we denote $laser[x]$ as the measured distance at angle $x$, including Front ($laser[0]$), Slightly Left, ($laser[45]$), Slightly Right

($laser[315]$), Left ($laser[90]$), and Right ($laser[270]$). If the obstacle distance in any of the five angles is less than the threshold, the proposed 'Obstacle-avoid Navigation' behavior is activated.

Algorithm 1 shows the pseudo-code for the 'Obstacle-avoid Navigation' behavior. The main idea is to identify the relative location of obstacles to the robot and command the robot to move around to avoid obstacles. Specifically, the robot initially set the linear velocity and angular velocity as $v_c$ and $\omega_c$, respectively. Positive values in $v_c$ and $\omega_c$ represent forward and left rotation, respectively, and negative values represent backward and right rotation, respectively. Initial values of $v_c$ and $\omega_c$ are set as 0.6 m/s and 0 rad/s in this work.
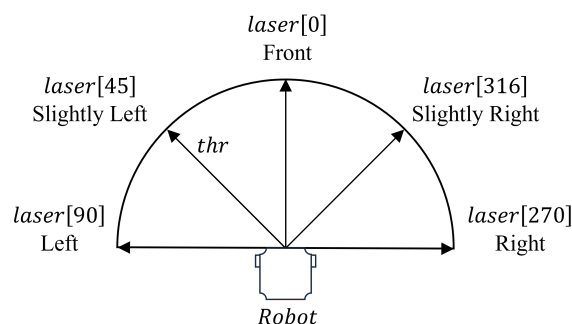


**Figure 3.** Five directions in the robot's laser distance sensing, including Left, Slightly Left, Front, Slightly Right, and Right. $laser[x]$ denotes the distance between the robot and the object at the angle $x$, which is measured from the onboard laser distance sensor.

---

**Algorithm 1** 'Obstacle-avoid Navigation' Behavior

---

1: Set robot linear velocity as $v_c = 0.6$ m/s
2: Set robot angular velocity as $\omega_c = 0$ rad/s
3: **if** $laser[0] > thr$ **then**
4:     $\omega_c = 0$ rad/s
5: **else**
6:     $v_c = 0$ m/s and $\omega_c = 0$ rad/s
7:     **if** $(laser[45] > thr) \vee (laser[315] > thr)$ **then**
8:         **if** $laser[45] > laser[315]$ **then**
9:             $\omega_c = 0.5$ rad/s
10:        **else**
11:            $\omega_c = -0.5$ rad/s
12:        **end if**
13:    **else if** $(laser[90] > thr) \vee (laser[270] > thr)$ **then**
14:        **if** $laser[90] > laser[270]$ **then**
15:            $\omega_c = 0.5$ rad/s
16:        **else**
17:            $\omega_c = -0.5$ rad/s
18:        **end if**
19:    **else**
20:        $v_c = -0.5$ m/s
21:    **end if**
22: **end if**

---

In the 'Obstacle-avoid Navigation' behavior, the robot will always check if there is a clear path in the Front direction, i.e., $laser[0] > thr$ ($thr$ is the threshold for obstacle detection, 0.75 m in this work), and if it is true, the robot will move forward with $\omega_c = 0$ rad/s. If the Front is blocked, the robot will stop and check Slightly Left or Slightly Right for a clear path ($(laser[45] > thr) \vee (laser[315] > thr)$). If there is a clear path in either of these two directions, the robot will compare clearance in Slightly Left and Slightly Right and rotate left (i.e., $\omega_c = 0.5$ rad/s) or right (i.e., $\omega_c = -0.5$ rad/s) to face the greater clearance. If there is no clearance in Slight Left or Slight Right, the robot will check Left and Right for a clear path ($(laser[90] > thr) \vee (laser[270] > thr)$). If there is a clear path, the robot

will compare Left and Right clearance ($laser[90] > laser[270]$) and rotate left ($\omega_c = 0.5$ rad/s) or right ($\omega_c = -0.5$ rad/s) to face the greater clearance. If there is no clear path in all five directions, the robot will move back ($v_c = -0.5$ m/s) to escape the dead end.

*3.4. Vision-based Navigation*

In this work, we employ vision as the main approach to detect odor sources within the search environment. Vision sensing allows the robot to detect the plume source location in its visual field and approach it directly. Olfaction-only navigation methods often rely on airflow direction for navigating to the odor source. This can lead to failure in turbulent airflow environments. Given visual sensing is not guided by airflow direction, combining it with Olfaction-based Navigation can allow the robot to find the odor source in turbulent airflow environments.

The proposed Vision-based Navigation relies on computer vision techniques. Specifically, we train a deep learning-based object detection model, i.e., YOLOv7, to detect vapors emitted from the odor source. Vapors can be considered as a common and distinct feature for the odor source object, such as smoke for fire sources, chemical plumes for chemical leaks or hydrothermal vents, etc. It should be mentioned that if the odor source does not have a distinct plume feature (i.e., transparent vapors), the robot can still find the odor source using the proposed Olfaction-based Navigation algorithm. We also provided real-world performance comparison between the Olfaction-based Navigation and the Vision and Olfaction Fusion Navigation algorithms.

In the proposed vision sensing method, we trained a YOLOv7 model to detect odor plumes in the continuously captured images. To generate training images, we extracted 243 observation frames with a resolution of $640 \times 480$ while the turtlebot was approaching the odor plumes in a variety of angles and lighting conditions. Figure 4 shows two sample frames used for training the vision model. This data was split into training, validation, and testing datasets for training the model. Roboflow [44] was utilized as the annotation tool for accurate bounding boxes and polygons delineation.
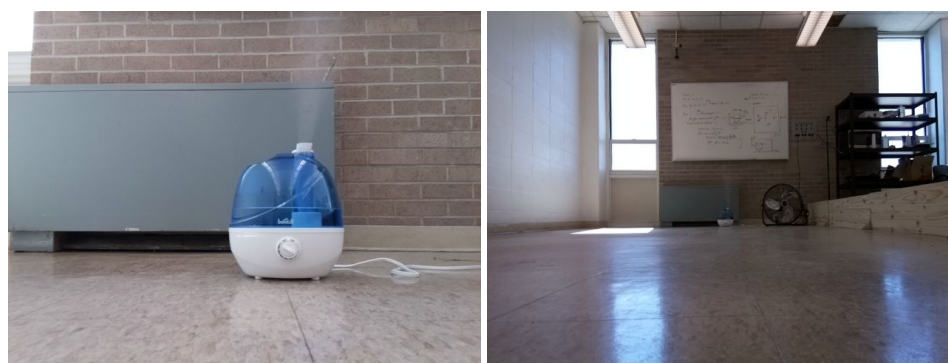


**Figure 4.** Two sample frames that include humidifier odor plumes in different lighting and spatial conditions. The frames are sampled out of the total 243 frames used for training the vision model. All of the frames were captured by the Turtlebot robot in the experiment area.

To assess YOLOv7 performance, diverse predefined augmentation techniques in Roboflow were systematically applied to 'Dataset-1'. These included rotation (-10° to +10°), shear (±15° horizontally and vertically), hue adjustment (-25° to +25°), saturation adjustment (-25% to +25%), brightness adjustment (-25% to +25%), exposure adjustment (-25% to +25%), blur (up to 2.5px), and noise (up to 1% of pixels). Post-augmentation, the resulting augmented dataset, labeled as 'Dataset-3', enriched the training set for a comprehensive evaluation of YOLOv7's robustness in detecting prescribed odor plumes. We set the number of training epochs to 100, with a batch size of 16. The resulting training accuracy was 98% and testing accuracy was 93%.

The implemented vision model returns a box bounding the plume in the image if it detects an odor plume. The output of the model also includes the horizontal and vertical location of the plume

bounding box. If the model returns a plume bounding box, the robot continues moving forward (i.e., $v_c = 0.5$ m/s) and checks if the horizontal location of the bounding box is in the left or the right half of the image. The model requires less than 1 second to generate output in our remote computer. The robot sends 30 image frames per second, and the robot picks every $30^{\text{th}}$ frame as the input to the vision model.

Equation 2 calculates robot's heading -

$$\omega_c = \begin{cases} 1 & 0.5 \text{ m/s if } c < \dfrac{w}{2} \\ 2 & -0.5 \text{ m/s if } c > \dfrac{w}{2}, \end{cases} \tag{2}$$

where $c$ is the horizontal mid-point of the bounding box, and w is the horizontal resolution of the captured image. If the bounding box is in the left half of the image (i.e., $c < \dfrac{w}{2}$), the robot rotates left (i.e., $\omega_c = 0.5$ rad/s) to face the plume. Otherwise, it rotates right ($\omega_c = -0.5$ rad/s) to face the plume.

*3.5. Olfaction-based Navigation*

If there is no valid visual detection but the robot can sense above-threshold odor concentration, Olfaction-based Navigation is employed to guide the robot to approach the odor source location.

Specifically, the proposed Olfaction-based Navigation algorithm commands the robot to move upwind to approach the odor source location. This behavior is analogous to the 'Surge' behavior of the bio-mimicking moth-inspired navigation OSL algorithm [45]. In this behavior, the robot's linear velocity is fixed at $v_c = 0.6$ m/s and the heading command, i.e., $\psi_c$, is calculated as:

$$\psi_c = \phi_{Inertial} + 180. \tag{3}$$

The robot will switch back to Vision-based Navigation once there is a valid vision detection.

*3.6. Source Declaration*

The robot is considered as successful if the robot position is within 0.9 m of the odor source location. But if the robot fails to reach the odor source within 200 seconds, the trial run is considered as a failure.

**4. Experiment Results**

*4.1. Search Area*

Figure 5 shows the 2-dimensional $8.2m \times 3.3m$ search area. Two obstacles were placed in the search area to simulate a complex real-world search environment. Ethanol vapor was used as the odor source as it is not toxic. Ethanol is also commonly implemented in OSL research [46]. A humidifier disperses ethanol vapor constantly as odor plume. To increase odor propagation in the search area, an electric fan was placed behind the humidifier. An additional fan was placed perpendicularly to the first fan to create a turbulent airflow environment. Using just Fan 1 creates a laminar airflow environment, and using both fans creates a turbulent airflow environment in the search area.
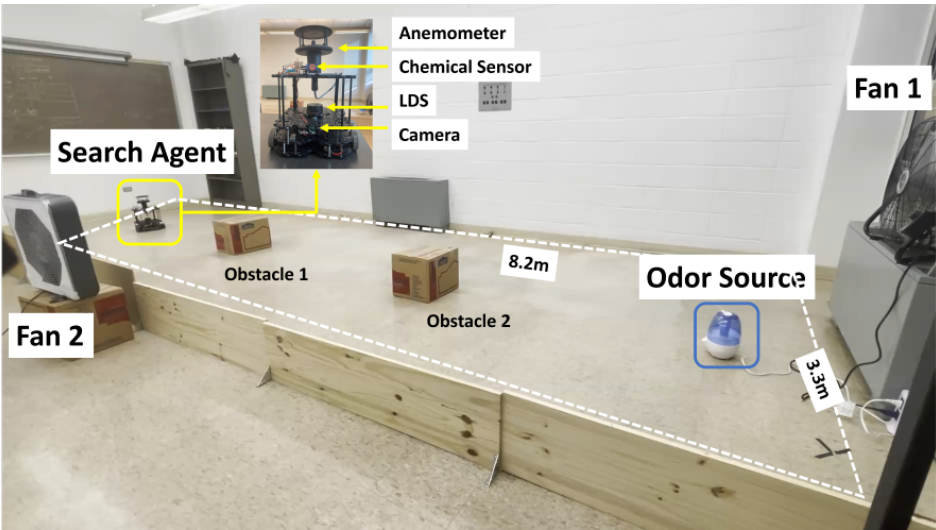
**Figure 5.** The experiment setup. The Turtlebot3 waffle pi mobile robot is used in this work. In addition to the camera and Laser Distance sensor, the robot is equipped with a chemical sensor and an anemometer for measuring chemical concentration, wind speeds, and directions. The robot is initially placed in a downwind area with the object of finding the odor source. A humidifier loaded with ethanol is employed to generate odor plumes. Two electric fans are placed perpendicularly to create artificial wind fields. Two obstacles are placed in the search area.

### 4.2. Mobile Robot Configuration

Turtlebot3 mobile robot platform was used in this work. Its built-in sensors include Raspberry Pi Camera, a 360-degree LiDAR sensor for sensing and a DYNAMIXEL diver for navigation. The onboard OpenCR controller allows the Turtlebot3 to be paired with additional sensors for increasing its functionalities.

Table 1 shows the built-in and added sensors for OSL experiments. Raspberry Pi Camera V2 was used for image capture, LDS-02 Laser Distance Sensor was used for obstacle detection, WindSonic Anemometer was used for wind speed and wind direction measurements in the body frame, and MQ3 alcohol detector was used for detecting chemical plume concentration.

**Table 1.** Type, name, and specification of the built-in camera, laser distance sensor, and added anemometer, chemical sensor.

| Source | Sensor Type | Module Name | Specification |
|---|---|---|---|
| **Built-in** | Camera | Raspberry pi camera v2 | Video Capture: 1080p30, 720p60 and VGA90. |
| | Laser Distance Sensor | LDS-02 | Detection Range: 360-degree. Distance Range: 160 ~8,000 mm. |
| **Added** | Anemometer | WindSonic, Gill Inc. | Speed: 0-75m/s. Wind direction: 0-360 degrees. |
| | Chemical Sensor | MQ3 alcohol detector | Concentration: 25 – 500 ppm. |

Turtlebot3 has Raspberry Pi 4 as the CPU which has limited computing power. It utilizes Ubuntu and Robot Operating System (ROS). Ubuntu allows connection capabilities with a remote computer. ROS allows custom programs in the remote computer to subscribe to specific sensor readings from the robot and publish heading commands back to the robot in real-time. ROS supports both Python and

C++ programming languages. Figure 6 presents the proposed system configuration for the robotic system, which includes a robotic agent, i.e., Turtlebot3, onboard controller, and a ground station, i.e., a remote Personal Computer (PC). For this study, Ubuntu 20.04 and ROS Noetic were installed in both the robot and the paired remote computer for controlling the robot. A local area network was used to connect the robot to the remote PC.
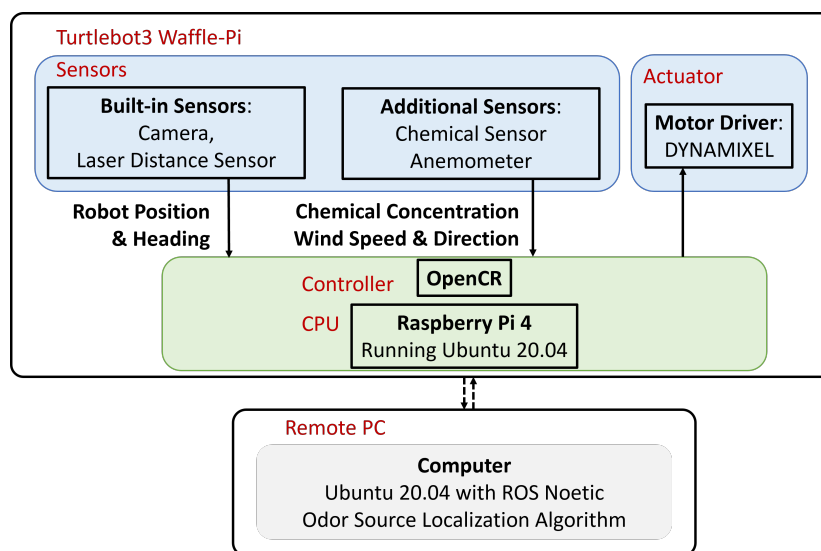


**Figure 6.** System configuration. This system contains two main components, including the Turtlebot3 and the remote PC. The solid connection line represents physical connection, and the dotted connection line represents wireless link.

### 4.3. Experiment Design

To determine the effectiveness of the proposed Vision and Olfaction Fusion Navigation algorithm, we tested the performance of Olfaction-only navigation and Vision-only navigation algorithms. Figure 7 shows the flow diagram of the two navigation algorithms. In the Olfaction-only navigation algorithm, the robot used the Crosswind maneuver behavior (Section 3.2), Obstacle-avoid Navigation behavior (Section 3.3), and Olfaction-based Navigation behavior (Section 3.5). In the absence of sufficient chemical concentration, the robot followed Crosswind maneuver behavior to maximize the chance of detecting sufficient plume concentration. If there were obstacles in the robot's path, it follows Obstacle-avoid Navigation behavior to circumvent the obstacles. If sufficient odor concentration is detected, and there are no obstacles in the robot's path, it follows Olfaction-based Navigation behavior to reach the odor source.

In the Vision-only navigation algorithm, the robot used the Crosswind maneuver behavior (Section 3.2), Obstacle-avoid Navigation behavior (Section 3.3), and Vision-based Navigation behavior (Section 3.4). In the absence of valid plume vision, the robot followed Crosswind maneuver behavior to maximize the chance of detecting valid plume vision. If there were obstacles in the robot's path, it follows Obstacle-avoid Navigation behavior to circumvent the obstacles. If the robot detects a valid plume visual, and there are no obstacles in the robot's path, it follows Vision-based Navigation behavior to reach the odor source.

These three algorithms were tested in two airflow environments, including the e1 - laminar airflow environment that used one electric fan and the e2 - turbulent airflow environment that used two perpendicularly placed electric fans. Thus, a total of six experiments setups were designed, i.e., three navigation methods in two airflow environments, to test the effectiveness of the proposed fusion model. Five experiment runs were conducted for each of the six experiment setups, totaling 30 trial runs. We used the same five starting positions to initialize the test runs. Figure 8 shows the five starting positions and the two airflow setups for the experiment runs.
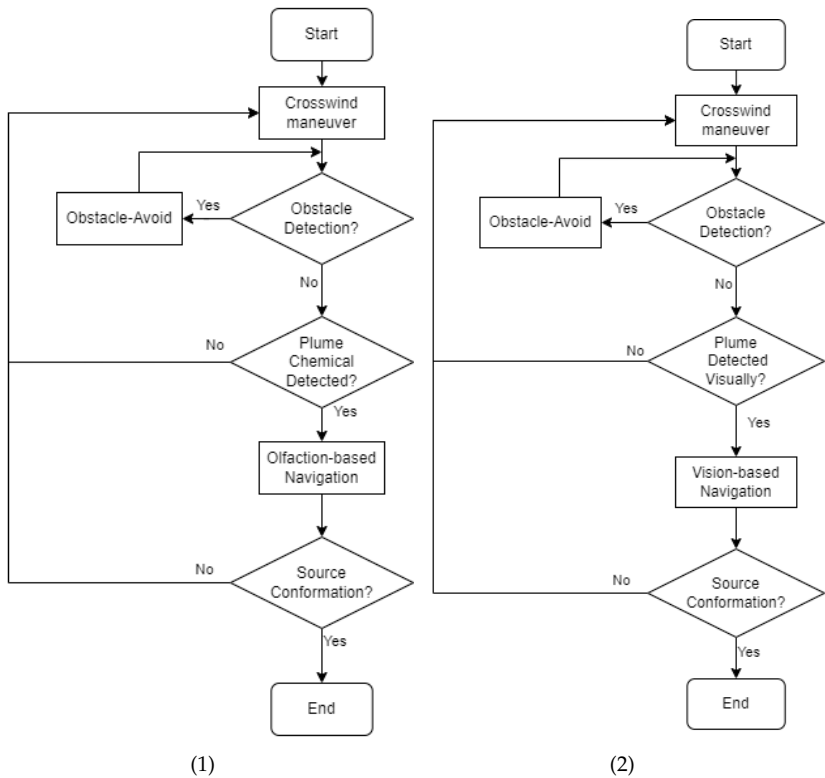
**Figure 7.** (1) The flow diagram of the Olfaction-only navigation algorithm. There are three navigation behaviors, including 'Crosswind maneuver', 'Obstacle-avoid Navigation', and 'Olfaction-based Navigation'. (2) The flow diagram of the Vision-only navigation algorithm. There are three navigation behaviors, including 'Crosswind maneuver', 'Obstacle-avoid Navigation', and 'Vision-based Navigation'.
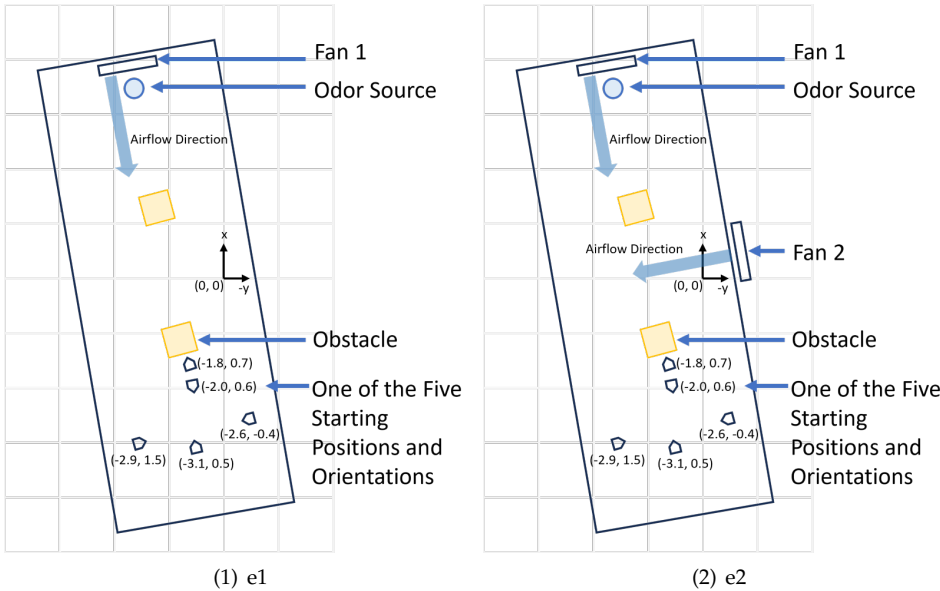


**Figure 8.** (1) The schematic diagram of the search area with e1 - laminar airflow setup. The five robot starting positions are used for testing the performance of the Olfaction-based Navigation, Vision-based Navigation, and Vision and Olfaction Fusion Navigation tests. (2) The schematic diagram of the search area with e2 - turbulent airflow setup.

*4.4. Sample Trials*

Figure 9 shows the robot trajectory and snapshots of the Vision and Olfaction Fusion Navigation trial run in a turbulent airflow environment. In this run, the robot initialized at $t$=1 s, found sufficient chemical concentration, and started following Olfaction-based Navigation. At $t$=22 s, the robot detected valid visual detection of the odor plumes and started to follow Vision-based Navigation. At $t$=49 s the robot faced the second obstacle and started to follow Obstacle-avoid Navigation behavior. It avoided the obstacle, re-detected plume vision, and started to follow Vision-based Navigation until it reached the odor source at $t$=72 s.
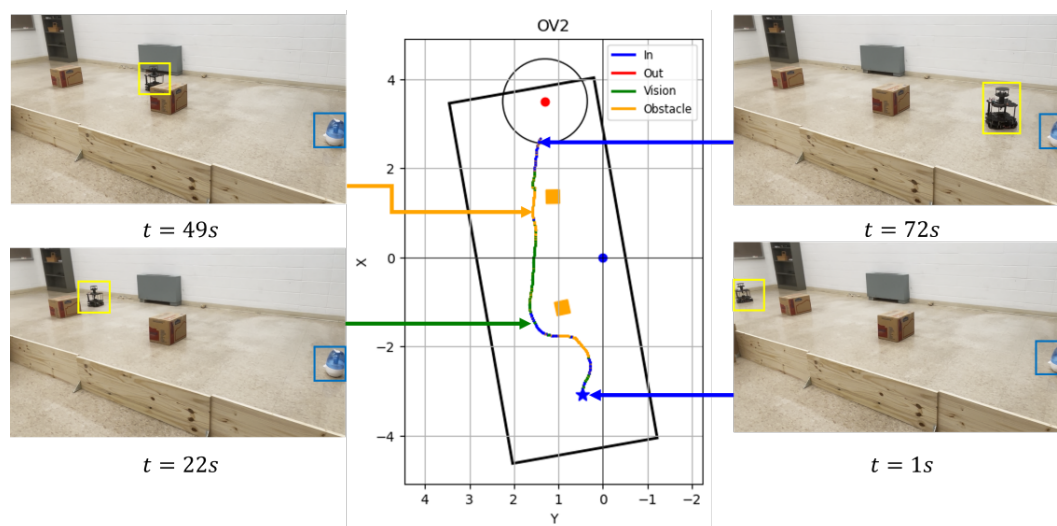


**Figure 9.** Robot trajectory graphs and snapshots of OSL tests with the Vision and Olfaction Fusion Navigation algorithm in turbulent airflow environment.

*4.5. Experiment Trials*

Table 2 shows the run times of the 30 trial runs, i.e., five trial runs using one of three navigation algorithms in two airflow environments. Figure 10 shows the robot trajectories in those 30 trial runs. Olfaction-only navigation algorithm uses airflow direction to navigate toward the odor source. It performed well in laminar airflow environments - the robot followed relatively direct airflow towards the odor source. However, in turbulent airflow environments, the robot got diverted by the complex airflow directions and often failed to reach the odor source by the designated time limit. Vision-based Navigation performed poorly in both laminar and turbulent airflow environments. Because of the obstacle placement, the robot had no visual of the plume from the starting position. It needed to follow the Crosswind maneuver and Obstacle-avoid Navigation behaviors until it had a valid plume vision. In most runs, the robot's 200-second time limit was over before it could find and navigate to the odor source. Vision and Olfaction Fusion Navigation algorithm test runs were consistently successful in both laminar and turbulent airflow environments. The Crosswind maneuver and Olfaction-based Navigation led the robot toward the odor source which allowed the robot to detect plume vision. Once it started to follow Vision-based Navigation, the robot was not affected by turbulent airflow.
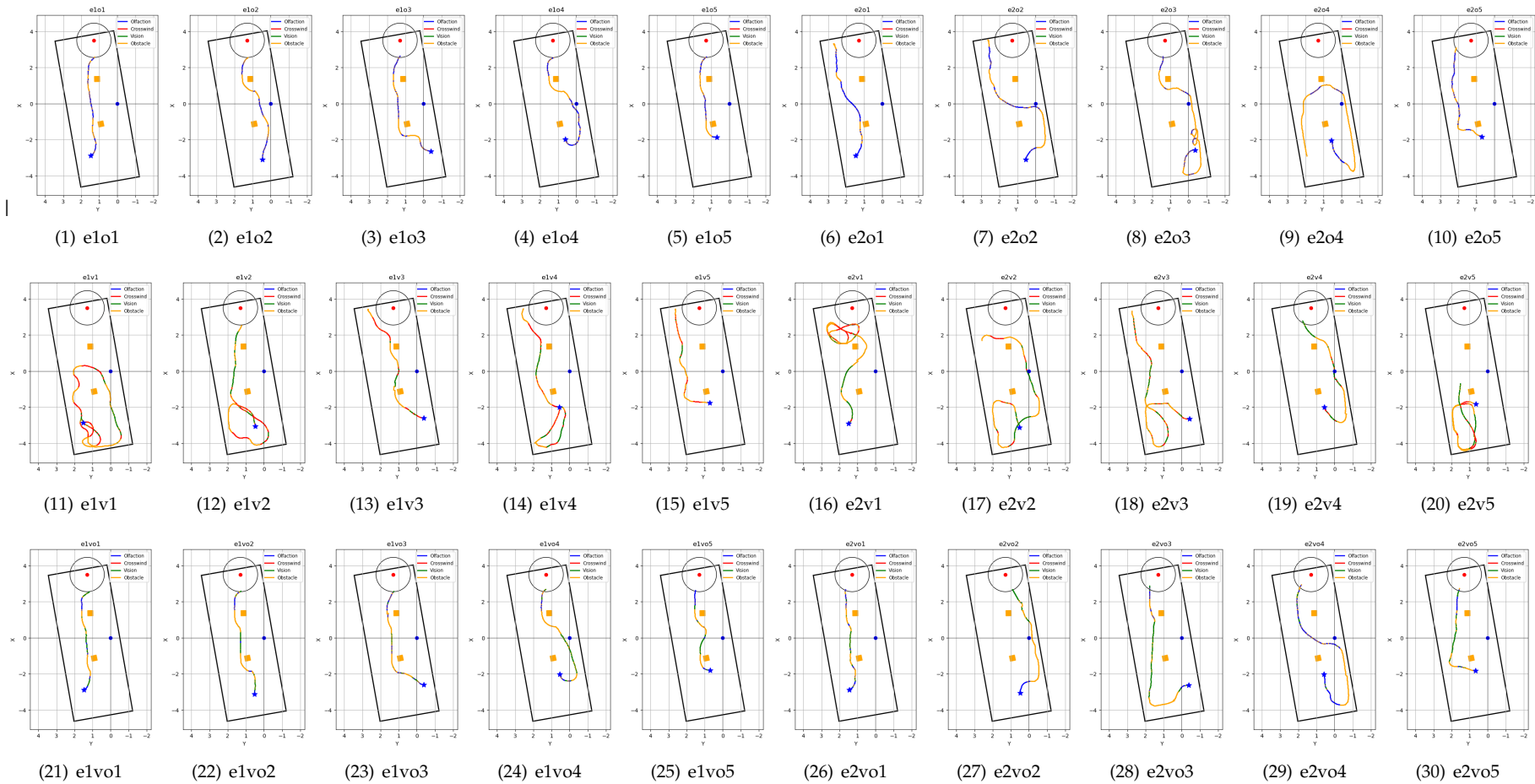
(1) e1o1     (2) e1o2     (3) e1o3     (4) e1o4     (5) e1o5     (6) e2o1     (7) e2o2     (8) e2o3     (9) e2o4     (10) e2o5

(11) e1v1    (12) e1v2    (13) e1v3    (14) e1v4    (15) e1v5    (16) e2v1    (17) e2v2    (18) e2v3    (19) e2v4    (20) e2v5

(21) e1vo1   (22) e1vo2   (23) e1vo3   (24) e1vo4   (25) e1vo5   (26) e2vo1   (27) e2vo2   (28) e2vo3   (29) e2vo4   (30) e2vo5

**Figure 10.** Trajectories of OSL repeat experiments. Olfaction-only Navigation algorithm trials (o1-o5) in - (1-5) laminar airflow environment (e1), and (6-10) turbulent airflow environment (e2). Similarly, Vision-only Navigation algorithm trials (v1-v5) in e1 (11-15) and e2 (16-20), Vision and Olfaction Fusion Navigation algorithm trials (vo1-vo5) in - e1 (21-25) and e2 (26-30). The behaviors that the robot was following under the three navigation algorithms are Crosswind - crosswind maneuver behavior, Obstacle - Obstacle-avoid Navigation behavior, Olfaction - Olfaction-based Navigation behavior, and Vision - Vision-based Navigation behavior. Robot starting positions are highlighted with a blue star, the obstacles are the orange boxes, and the odor source is the red point with the surrounding circular source declaration region.

**Table 2.** Search Time of the Vision-only, Olfaction-only, and the Proposed Vision and Olfaction Fusion Navigation Algorithms.

| | Robot Initial Position (x, y), Orientation (z, w) | Olfaction-only Navigation Algorithm (s) | Vision-only Navigation Algorithm (s) | Vision and Olfaction Fusion Navigation Algorithm (s) |
|---|---|---|---|---|
| **Laminar Airflow Env.** | (-2.9, 1.5), (-0.6, 1.0) | **63.1** | - | 63.9 |
| | (-3.1, 0.5), (0.0, 35.0) | 71.3 | 149.3 | **69.9** |
| | (-2.6, -0.4), (0.7, 0.7) | 74.3 | - | **67.5** |
| | (-2.0, 0.6), (1.0, -0.1) | **73.8** | - | 75.7 |
| | (-1.8, 0.7), (0.0, 0.1) | **59.1** | - | 61.1 |
| **Turbulent Airflow Env.** | (-2.9, 1.5), (-0.6, 1.0) | - | - | **64.0** |
| | (-3.1, 0.5), (0.0, 35.0) | - | - | **113.1** |
| | (-2.6, -0.4), (0.7, 0.7) | 196.4 | - | **130.7** |
| | (-2.0, 0.6), (1.0, -0.1) | - | 102.8 | **131.9** |
| | (-1.8, 0.7), (0.0, 0.1) | 72.3 | - | **68.5** |

*4.6. Statistic Analysis*

Figure 11 shows the combined robot trajectories of the three navigation algorithms in the two airflow environments. Table 3 summarizes the repeated test results in terms of success rate, averaged search time, and average traveled distance. For failed experiment runs, 200 s was used for calculating the Average Search Time (s). We can observe from the results that the proposed Vision and Olfaction Fusion Navigation algorithm has the highest success rate, the lowest average search time, and the lowest average distance traveled among the three methods. This is critical in real-world odor source localization applications, as we want the robot to find odor sources as quickly as possible.

**Table 3.** Result Statistics, i.e., Success Rate and Average Search Time of Vision-based Navigation, Olfaction-based Navigation, and the Proposed Vision and Olfaction Fusion Navigation Algorithms.

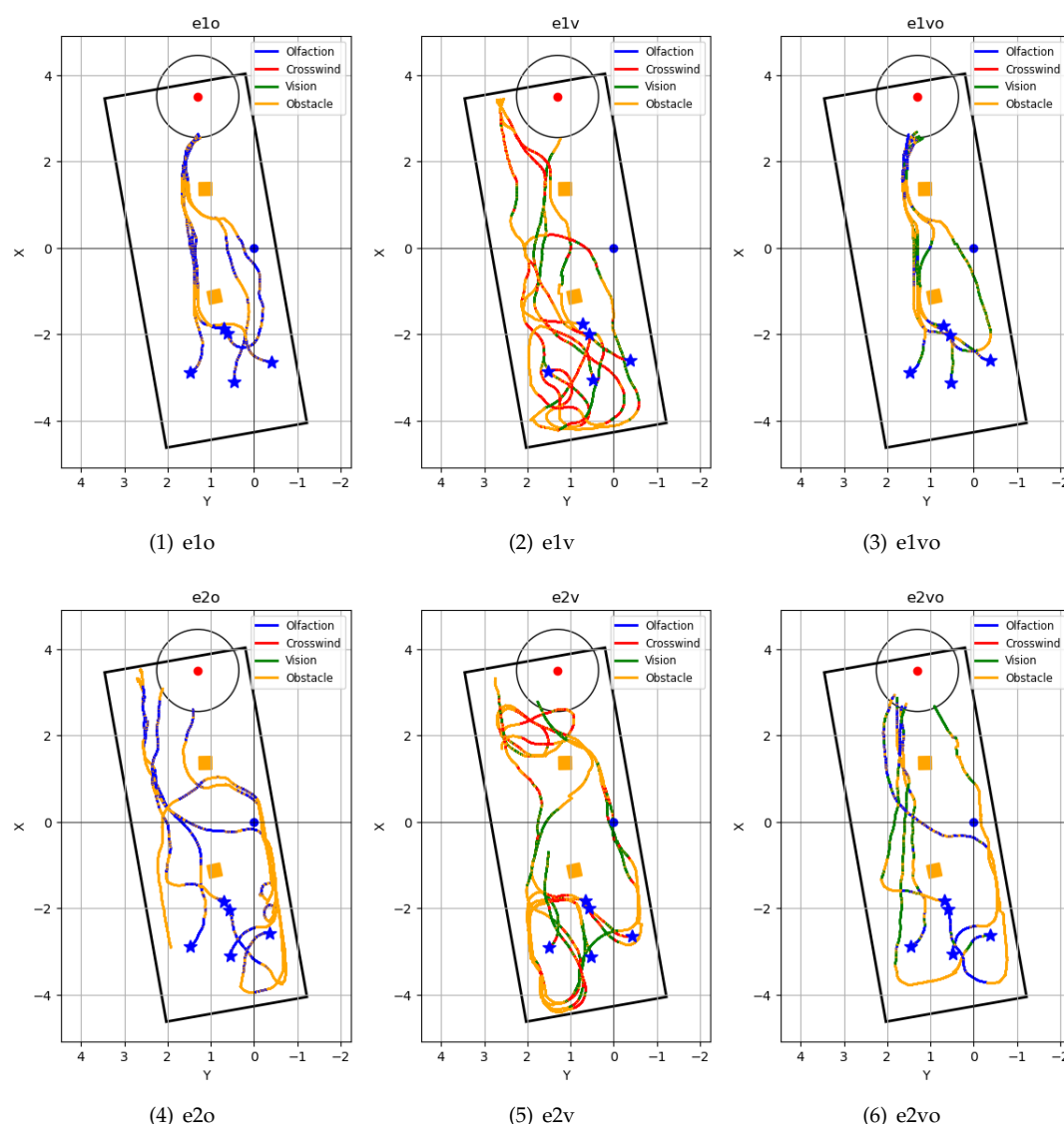| Airflow Environment | Navigation Algorithm | Success Rate | Avg. Search Time (s) | Avg. Travelled Distance (m) |
|---|---|---|---|---|
| **Laminar** | **Olfaction-only** | 5/5 | 68.3 | **6.1** |
| | **Vision-only** | 1/5 | 189.9 | 11.7 |
| | **Vision and Olfaction Fusion** | 5/5 | **67.6** | 6.2 |
| **Turbulent** | **Olfaction-only** | 2/5 | 173.7 | 9.7 |
| | **Vision-only** | 1/5 | 180.6 | 13.7 |
| | **Vision and Olfaction Fusion** | 5/5 | **101.6** | 7.8 |
| **Combined** | **Olfaction-only** | 7/10 | 121.0 | 7.9 |
| | **Vision-only** | 2/10 | 185.2 | 12.7 |
| | **Vision-Olfaction Fusion** | **10/10** | **84.6** | **7.0** |

**Figure 11.** Robot trajectories of repeated tests in six navigation algorithm and airflow environment combinations. Trajectories in laminar airflow environments are - (1) e1o - Olfaction-only navigation algorithm, (2) e1v - Vision-only navigation algorithm, and (3) e1vo - Vision and Olfaction Fusion Navigation algorithm. Trajectories in turbulent airflow environment are - (4) e2o - Olfaction-only navigation algorithm, (5) e2v - Vision-only navigation algorithm, (6) e2vo - Vision and Olfaction Fusion Navigation algorithm. The behaviors that the robot was following under the three navigation algorithms are shown in the trajectory. These behaviors include Crosswind - crosswind maneuver behavior, Obstacle - Obstacle-avoid Navigation behavior, Olfaction - Olfaction-based Navigation behavior, and Vision - Vision-based Navigation behavior. Five robot starting positions are highlighted with a blue star, the obstacles are the orange boxes, and the odor source is the red point with the surrounding circular source declaration region.

## 5. Future Research Direction

A number of improvements can be made to the proposed OSL algorithm in the future. Firstly, the proposed navigation algorithm follows a homogeneous crosswind maneuver behavior for finding odor plumes. The search behavior doesn't take into account past vision or olfaction sensing history. Similarly, the moth-inspired algorithm used in this paper only uses current olfaction readings for finding the odor source. Whereas engineering-based solutions like the Particle Filter utilize past sensor readings for estimating the odor source and plume location. Thus, future research scope includes

pairing engineering-based Olfaction navigation with Vision-based Navigation for improved crosswind maneuver and Olfaction-based Navigation. The implemented Obstacle-avoid Navigation algorithm in this paper also relies only on the current laser readings to sense and circumvent obstacles. In this case, reactive path planning algorithms include Fuzzy Logic, Neural Networks, bug algorithms, etc. [47] can be adopted for more efficient Obstacle-avoid Navigation behavior. Additionally, the future scope of this robot platform includes using machine learning algorithms for calculating robot headings. For instance, the reinforcement learning (RL) [48] and supervised learning [49] methods can be used for olfactory-based navigation in robots. Transformer-based Vision-Language and Vision-Language-Action (VLA) models are gaining traction as a prevalent approach in robotics. Recent applications of such a model include the PaLM-E model [50], and the RT-2 [51]. Exploring the possibilities of the Vision-Language models as the primary decision-maker for multi-modal odor source localization is another exciting possibility in OSL research.

## 6. Conclusion

The combination of computer vision and robotic olfaction provides a more comprehensive observation of the environment, enabling the robot to interact with the environment in more ways and enhancing navigation performance. This paper proposes the incorporation of vision sensing in OSL. Specifically, the paper proposes a Vision and Olfaction Fusion Navigation algorithm with Obstacle-avoid Navigation capability for 2-D odor source localization tasks for ground mobile robots. For conducting real-world experiments to test the proposed algorithm, a robot platform based on the Turtlebot3 mobile robot was developed with olfaction and vision-sensing capabilities. The proposed navigation algorithm had five behaviors, i.e., Crosswind maneuver behavior to find odor plume, Obstacle-avoid Navigation behavior to circumvent obstacles in the environment, Vision-based Navigation to approach the odor source using vision sensing, Olfaction-based Navigation to approach the odor source using olfaction sensing, and source declaration. For the Vision-based Navigation behavior, a YOLOv7-based vision model was trained to detect visible odor plumes. For Olfaction-based Navigation behavior, we used moth-inspired algorithm. To test the performance of the proposed Vision and Olfaction Fusion Navigation algorithm, we tested the performance of the Olfaction-only navigation algorithm, Vision-only navigation algorithm, and the proposed Vision and Olfaction Fusion Navigation algorithm separately in real-world experiment setups. Furthermore, we tested the performance of the three navigation algorithms in laminar and turbulent airflow environments to compare their strengths. We used five predefined starting robot positions for each navigation algorithm and repeated them for both airflow environments - resulting in 30 total experiment runs. The search results of the OSL experiments show that the proposed Vision and Olfaction Fusion Navigation algorithm had a higher success rate, lower average search time, and lower average traveled distance for finding the odor source compared to Olfaction-only and Vision-only navigation algorithms in both laminar and turbulent airflow environments. The result of our experiment indicates that vision sensing is a promising addition to olfaction sensing in ground-mobile robot-based Odor Source Localization research.

**Conflicts of Interest:** The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

## Abbreviations

The following abbreviations are used in this manuscript:

| | |
|---|---|
| AUV | Autonomous Underwater Vehicle |
| ANFIS | Adaptive Neuro-fuzzy Inference System |
| DL | Deep Learning |
| DNN | Deep Neural Networks |
| LDS | Laser Distance Sensor |
| OSL | Odor Source Localization |
| PC | Personal Computer |
| ROS | Robot Operating System |
| SLAM | Simultaneous Localization and Mapping |
| VLA | Vision-Language-Action model |

## References

1. Purves, D.; Augustine, G.; Fitzpatrick, D.; Katz, L.; LaMantia, A.; McNamara, J.; Williams, S. The Organization of the Olfactory System. *Neuroscience* **2001**, pp. 337–354.
2. Sarafoleanu, C.; Mella, C.; Georgescu, M.; Perederco, C. The importance of the olfactory sense in the human behavior and evolution. *Journal of Medicine and life* **2009**, *2*, 196.
3. Ibarra-Soria, X.; Levitin, M.O.; Saraiva, L.R.; Logan, D.W. The olfactory transcriptomes of mice. *PLoS genetics* **2014**, *10*, e1004593.
4. Kowadlo, G.; Russell, R.A. Robot odor localization: a taxonomy and survey. *The International Journal of Robotics Research* **2008**, *27*, 869–894.
5. Wang, L.; Pang, S.; Noyela, M.; Adkins, K.; Sun, L.; El-Sayed, M. Vision and Olfactory-Based Wildfire Monitoring with Uncrewed Aircraft Systems. 2023 20th International Conference on Ubiquitous Robots (UR). IEEE, 2023, pp. 716–723.
6. Fu, Z.; Chen, Y.; Ding, Y.; He, D. Pollution source localization based on multi-UAV cooperative communication. *IEEE Access* **2019**, *7*, 29304–29312.
7. Burgués, J.; Hernández, V.; Lilienthal, A.J.; Marco, S. Smelling nano aerial vehicle for gas source localization and mapping. *Sensors* **2019**, *19*, 478.
8. Russell, R.A. Robotic location of underground chemical sources. *Robotica* **2004**, *22*, 109–115.
9. Chen, Z.; Wang, J. Underground odor source localization based on a variation of lower organism search behavior. *IEEE Sensors Journal* **2017**, *17*, 5963–5970.
10. Wang, L.; Pang, S.; Xu, G. 3-dimensional hydrothermal vent localization based on chemical plume tracing. Global Oceans 2020: Singapore–US Gulf Coast. IEEE, 2020, pp. 1–7.
11. Cardé, R.T.; Mafra-Neto, A. Mechanisms of flight of male moths to pheromone. In *Insect pheromone research*; Springer, 1997; pp. 275–290.
12. López, L.L.; Vouloutsi, V.; Chimeno, A.E.; Marcos, E.; i Badia, S.B.; Mathews, Z.; Verschure, P.F.; Ziyatdinov, A.; i Lluna, A.P. Moth-like chemo-source localization and classification on an indoor autonomous robot. In *On Biomimetics*; IntechOpen, 2011. doi:10.5772/19695.
13. Zhu, H.; Wang, Y.; Du, C.; Zhang, Q.; Wang, W. A novel odor source localization system based on particle filtering and information entropy. *Robotics and autonomous systems* **2020**, *132*, 103619.
14. Jing, T.; Meng, Q.H.; Ishida, H. Recent progress and trend of robot odor source localization. *IEEJ Transactions on Electrical and Electronic Engineering* **2021**, *16*, 938–953.
15. Berg, H.C. Feature article site index motile behavior of bacteria. *Physics today* **2001**, *9*, 25.
16. Lockery, S.R. The computational worm: spatial orientation and its neuronal basis in C. elegans. *Current opinion in neurobiology* **2011**, *21*, 782–790.
17. Radvansky, B.A.; Dombeck, D.A. An olfactory virtual reality system for mice. *Nature communications* **2018**, *9*, 839.

18. Sandini, G.; Lucarini, G.; Varoli, M. Gradient driven self-organizing systems. Proceedings of 1993 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'93). IEEE, 1993, Vol. 1, pp. 429–432.

19. Grasso, F.W.; Consi, T.R.; Mountain, D.C.; Atema, J. Biomimetic robot lobster performs chemo-orientation in turbulence using a pair of spatially separated sensors: Progress and challenges. *Robotics and Autonomous Systems* **2000**, *30*, 115–131.

20. Russell, R.A.; Bab-Hadiashar, A.; Shepherd, R.L.; Wallace, G.G. A comparison of reactive robot chemotaxis algorithms. *Robotics and Autonomous Systems* **2003**, *45*, 83–97.

21. Lilienthal, A.; Duckett, T. Experimental analysis of gas-sensitive Braitenberg vehicles. *Advanced Robotics* **2004**, *18*, 817–834.

22. Ishida, H.; Nakayama, G.; Nakamoto, T.; Moriizumi, T. Controlling a gas/odor plume-tracking robot based on transient responses of gas sensors. *IEEE Sensors Journal* **2005**, *5*, 537–545.

23. Murlis, J.; Elkinton, J.S.; Carde, R.T. Odor plumes and how insects use them. *Annual review of entomology* **1992**, *37*, 505–532.

24. Vickers, N.J. Mechanisms of animal navigation in odor plumes. *The Biological Bulletin* **2000**, *198*, 203–212.

25. Cardé, R.T.; Willis, M.A. Navigational strategies used by insects to find distant, wind-borne sources of odor. *Journal of chemical ecology* **2008**, *34*, 854–866.

26. Nevitt, G.A. Olfactory foraging by Antarctic procellariiform seabirds: life at high Reynolds numbers. *The Biological Bulletin* **2000**, *198*, 245–253.

27. Wallraff, H.G. Avian olfactory navigation: its empirical foundation and conceptual state. *Animal Behaviour* **2004**, *67*, 189–204.

28. Shigaki, S.; Sakurai, T.; Ando, N.; Kurabayashi, D.; Kanzaki, R. Time-varying moth-inspired algorithm for chemical plume tracing in turbulent environment. *IEEE Robotics and Automation Letters* **2017**, *3*, 76–83.

29. Shigaki, S.; Shiota, Y.; Kurabayashi, D.; Kanzaki, R. Modeling of the Adaptive Chemical Plume Tracing Algorithm of an Insect Using Fuzzy Inference. *IEEE Transactions on Fuzzy Systems* **2019**, *28*, 72–84. doi:10.1109/tfuzz.2019.2915187.

30. Rahbar, F.; Marjovi, A.; Kibleur, P.; Martinoli, A. A 3-D bio-inspired odor source localization and its validation in realistic environmental conditions. 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2017, pp. 3983–3989.

31. Vergassola, M.; Villermaux, E.; Shraiman, B.I. 'Infotaxis' as a strategy for searching without gradients. *Nature* **2007**, *445*, 406.

32. Jakuba, M.V. Stochastic mapping for chemical plume source localization with application to autonomous hydrothermal vent discovery. PhD thesis, Massachusetts Institute of Technology, 2007. doi:10.1575/1912/1583.

33. Rahbar, F.; Marjovi, A.; Martinoli, A. An algorithm for odor source localization based on source term estimation. 2019 International Conference on Robotics and Automation (ICRA). IEEE, 2019, pp. 973–979.

34. Hutchinson, M.; Liu, C.; Chen, W.H. Information-based search for an atmospheric release using a mobile robot: Algorithm and experiments. *IEEE Transactions on Control Systems Technology* **2018**, *27*, 2388–2402. doi:10.1109/TCST.2018.2860548.

35. Jiu, H.; Chen, Y.; Deng, W.; Pang, S. Underwater chemical plume tracing based on partially observable Markov decision process. *International Journal of Advanced Robotic Systems* **2019**, *16*, 1729881419831874.

36. Pang, S.; Zhu, F. Reactive planning for olfactory-based mobile robots. 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems. IEEE, 2009, pp. 4375–4380.

37. Wang, L.; Pang, S. Chemical Plume Tracing using an AUV based on POMDP Source Mapping and A-star Path Planning. OCEANS 2019 MTS/IEEE SEATTLE. IEEE, 2019, pp. 1–7.

38. Kim, H.; Park, M.; Kim, C.W.; Shin, D. Source localization for hazardous material release in an outdoor chemical plant via a combination of LSTM-RNN and CFD simulation. *Computers & Chemical Engineering* **2019**, *125*, 476–489.

39. Hu, H.; Song, S.; Chen, C.P. Plume Tracing via Model-Free Reinforcement Learning Method. *IEEE transactions on neural networks and learning systems* **2019**.

40. Wang, L.; Pang, S. An Implementation of the Adaptive Neuro-Fuzzy Inference System (ANFIS) for Odor Source Localization. 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems. IEEE, 2021.

41. Baker, K.L.; Dickinson, M.; Findley, T.M.; Gire, D.H.; Louis, M.; Suver, M.P.; Verhagen, J.V.; Nagel, K.I.; Smear, M.C. Algorithms for olfactory search across species. *Journal of Neuroscience* **2018**, *38*, 9383–9389.

42. Liu, A.; Papale, A.E.; Hengenius, J.; Patel, K.; Ermentrout, B.; Urban, N.N. Mouse navigation strategies for odor source localization. *Frontiers in Neuroscience* **2020**, *14*, 218.

43. Li, W.; Farrell, J.A.; Pang, S.; Arrieta, R.M. Moth-inspired chemical plume tracing on an autonomous underwater vehicle. *IEEE Transactions on Robotics* **2006**, *22*, 292–307.

44. Ciaglia, F.; Zuppichini, F.S.; Guerrie, P.; McQuade, M.; Solawetz, J. Roboflow 100: A Rich, Multi-Domain Object Detection Benchmark. *arXiv preprint arXiv:2211.13523* **2022**.

45. Farrell, J.A.; Pang, S.; Li, W. Chemical plume tracing via an autonomous underwater vehicle. *IEEE Journal of Oceanic Engineering* **2005**, *30*, 428–442.

46. Feng, Q.; Cai, H.; Chen, Z.; Yang, Y.; Lu, J.; Li, F.; Xu, J.; Li, X. Experimental study on a comprehensive particle swarm optimization method for locating contaminant sources in dynamic indoor environments with mechanical ventilation. *Energy and buildings* **2019**, *196*, 145–156.

47. Patle, B.; Pandey, A.; Parhi, D.; Jagadeesh, A.; others. A review: On path planning strategies for navigation of mobile robot. *Defence Technology* **2019**, *15*, 582–606.

48. Wang, L.; Pang, S.; Li, J. Olfactory-Based Navigation via Model-Based Reinforcement Learning and Fuzzy Inference Methods. *IEEE Transactions on Fuzzy Systems* **2021**, *29*, 3014–3027. doi:10.1109/TFUZZ.2020.3011741.

49. Wang, L.; Yin, Z.; Pang, S. Learn to Trace Odors: Robotic Odor Source Localization via Deep Learning Methods with Real-world Experiments. SoutheastCon 2023. IEEE, 2023, pp. 524–531.

50. Driess, D.; Xia, F.; Sajjadi, M.S.; Lynch, C.; Chowdhery, A.; Ichter, B.; Wahid, A.; Tompson, J.; Vuong, Q.; Yu, T.; others. Palm-e: An embodied multimodal language model. *arXiv preprint arXiv:2303.03378* **2023**.

51. Brohan, A.; Brown, N.; Carbajal, J.; Chebotar, Y.; Chen, X.; Choromanski, K.; Ding, T.; Driess, D.; Dubey, A.; Finn, C.; others. Rt-2: Vision-language-action models transfer web knowledge to robotic control. *arXiv preprint arXiv:2307.15818* **2023**.