Semantic Odor Source Localization via Visual and Olfactory Integrated Navigation

1st Lingxiao Wang
Dept. of Electrical Engineering
Louisiana Tech University
Ruston, USA
lwang@latech.edu

2nd Sunzid Hassan
Dept. of Computer Science
Louisiana Tech University
Ruston, USA
sha040@latech.edu

3rd Khan Raqib Mahmud
Dept. of Computer Science
Louisiana Tech University
Ruston, USA
krm070@latech.edu

***Abstract*—Odor Source Localization (OSL) is a technology that navigates a mobile robot to autonomously locate a hidden odor source. Unlike traditional OSL navigation algorithms, which rely solely on olfactory data, this paper introduces a semantic OSL navigation algorithm that integrates both visual and olfactory sensing to enhance the search performance. By combining these two modalities, the proposed system can infer potential odor sources and their locations. For example, when detecting the smell of smoke in a kitchen, our system can associate the odor source with an oven or microwave. To leverage the semantic relationships between visual and olfactory observations, we employ a Large Language Model (LLM) to process the multi-modal sensory data and guide the navigation. The proposed LLM-based navigation algorithm is evaluated in a simulated household environment. Simulated results demonstrate that the proposed method can achieve a higher success rate and shorter travel distance, compared to random walk, vision-only, and olfaction-only approaches.**

***Index Terms*—Robotic Odor Source Localization, large language models, robot navigation, household robots.**

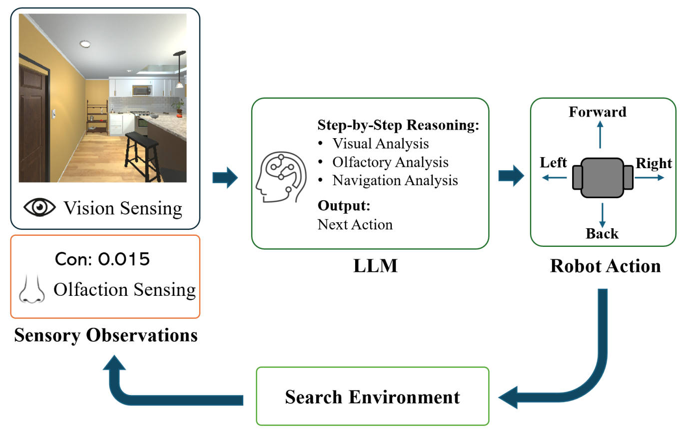

Fig. 1. An overview of the proposed LLM-based semantic OSL navigation algorithm. The LLM will be provided with both vision and olfaction sensory information to infer possible odor source objects and generate an action to guide the robot toward the odor source.

I. Introduction

Robotic Odor Source Localization (OSL) enables a mobile robot to autonomously locate hidden odor sources [1]–[3]. This technology can replace humans in hazardous environments and has a wide range of real-world applications, including locating chemical gas leaks [4], detecting hydrothermal vents [5], identifying wildfire locations [6], [7], and monitoring air pollution [8]. The effectiveness of OSL depends on the design of a robust navigation algorithm, which directs the robot toward the odor source based on environmental observations. Similar to image-based navigation algorithms [9], which use visual cues to guide movement, traditional OSL navigation algorithms rely solely on olfactory data, such as gas or chemical concentration levels, to inform the search process.

This work proposes a semantic OSL navigation algorithm that integrates both visual and olfactory sensory information to improve search performance. The fusion of vision and olfaction is inspired by human odor-searching behavior [10]. For instance, when we detect the aroma of coffee, our instinct is to visually scan the environment for a possible source. If we spot a cup of dark liquid, we use vision to approach it and olfaction to confirm. When no clear visual target is available, we rely on olfactory cues to approach the odor source. In this process, the olfactory sensing data contains rich semantic information that can help infer likely odor source objects and locations. For example, upon detecting the smell of coffee, we might infer that the source is a coffee cup or kettle and that it is likely located on a table or in a kitchen.

Recent breakthroughs in Large Language Models (LLMs) have enabled artificial intelligence (AI) systems to interpret semantics in textual [11] and visual inputs [12]. Leveraging these advancements, our proposed OSL navigation agent utilizes LLMs to process both visual and olfactory sensing data. As illustrated in Fig. 1, the LLM receives visual inputs, i.e., the raw images captured from the robot's front-facing camera, alongside the olfactory data, i.e., the odor concentrations and the type of detected odors. Instead of relying on a pre-trained computer vision model (e.g., YOLO [13]), we directly input the captured images into the LLM. The model then determines the most suitable navigation action from a predefined action pool, guiding the robot toward the odor source. We summarize our contributions as follows:

This work is supported by Louisiana Board of Regents with the contract number: LEQSF(2024-27)-RD-A-22.

- We propose an LLM-based semantic OSL navigation algorithm. To the best of our knowledge, this is the first work to leverage semantic information from both visual and olfactory observations using LLMs for OSL tasks.
- We integrate Chain-of-Thought (CoT) reasoning [14] into the LLM decision-making process, guiding the robot to connect visual and olfactory sensing information.
- We evaluate the proposed algorithm in a simulated household environment, comparing its performance against random walks and sole-modality searching strategies. Results demonstrate that our semantic OSL navigation algorithm achieves higher success rates and shorter travel distances.

The remainder of this paper is structured as follows: Section II reviews related works. Section III presents the proposed semantic OSL algorithm. Section IV presents the simulated household environment and the algorithm's performance evaluation. Section V discusses conclusions and future research directions.

II. Related Works

A. Traditional OSL Navigation Algorithm

Traditional Odor Source Localization (OSL) algorithms can be categorized into three main groups: chemotaxis-based methods, bio-inspired methods, and engineering-based (or probabilistic) methods.

Chemotaxis-based methods guide the robot to trace odor plumes by following the odor concentration gradient [15]. A typical setup involves two chemical sensors placed on either side of the robot, allowing it to move toward the side with the higher detected concentration [16], [17]. This approach is effective in laminar flow environments, where odor plumes disperse in a steady and coherent manner. However, in turbulent flow environments, where plumes become patchy and intermittent, the concentration gradient is often unreliable, making chemotaxis significantly less effective.

Bio-inspired methods mimic the olfactory behaviors observed in animals. For instance, moth-inspired strategies employ a "surge/casting" behavior, where the robot surges upwind when detecting an odor and switches to casting across the wind when the plume is lost [5], [18], [19]. Another example is the lobster-inspired method, which enhances chemotaxis by turning toward the side with a higher odor concentration and retracing its path when both sensors detect equal concentrations [20]. Although bio-inspired methods are often simple and computationally efficient, in turbulent flow environments, the cross-wind casting strategy can be less effective, resulting in longer search times and even search failures [21].

Probabilistic methods use mathematical and physical models to estimate the distribution of odor plumes and predict the odor source's location. The search area is divided into cells, each with a probability of containing the source. As the robot explores, these probabilities are updated based on observations, eventually converging to a specific region that likely contains the odor source. Algorithms for calculating source probabilities include Bayesian inference [22], particle filters [23], and partially observable Markov decision processes (POMDPs) [24]. Path planning algorithms such as artificial potential fields (APF) [25] and A-star [26] are then used to direct the robot toward the estimated source. While probabilistic methods provide a systematic and data-driven approach to odor localization, they are computationally intensive, especially in large search areas with many cells [21], which can be a limitation for robots with constrained computational resources. Additionally, these methods often rely on simplified plume models and approximate global wind conditions using local measurements, which can introduce errors in source estimation [22], [23].

B. LLM-based Navigation

Multi-modal LLM-based robot navigation is an emerging paradigm that leverages human-like high-level semantic understanding and reasoning abilities to enhance autonomous decision-making. By integrating multi-modal LLMs with conventional navigation systems, these approaches can analyze human commands alongside sensor inputs to generate intelligent, context-aware strategies for maneuvering in unforeseen scenarios [27]. A research problem associated with LLM-based navigation is the Object Goal Navigation (ObjectNav), where the robotic agent needs to locate an object based on the textual description of this object. Several different approaches have been developed for this task, including CLIP on Wheels (CoW) [28], VLFM [29], CLIP-Nav [30], VELMA [31], LM-Nav [32], and A^2Nav [33]. A core setup of these approaches is to utilize the LLM as the "brain" of the robot to process visual observations and guide the robot moving toward the target object.

C. Research Niche

Compared to existing olfactory-based navigation algorithms and ObjectNav methods, our approach introduces a novel strategy by integrating two distinct modalities, i.e., vision and olfaction, to infer search target locations. We leverage LLMs to extract semantic information from visual observations (similar to ObjectNav methods) to identify potential odor source objects while using olfactory observations to guide the search direction and enhance efficiency (like olfactory-based navigation). By combining these two modalities, the robot can dynamically adapt to its environment, improving its ability to locate odor sources more effectively.

III. Methodology

A. Problem Formation

A ground mobile robot is employed as the search agent, and it is equipped with a camera for visual detection, a chemical sensor for olfactory detection, and a depth camera for obstacle detection. At each time step t during the search, the robot receives visual observations v and chemical concentration ρ at the robot's position. In a vector form, the sensory observation is denoted as: $\mathbf{o}^t = [v, \rho]^t$. The output of the navigation algorithm is an action a^t that directs the robot to move

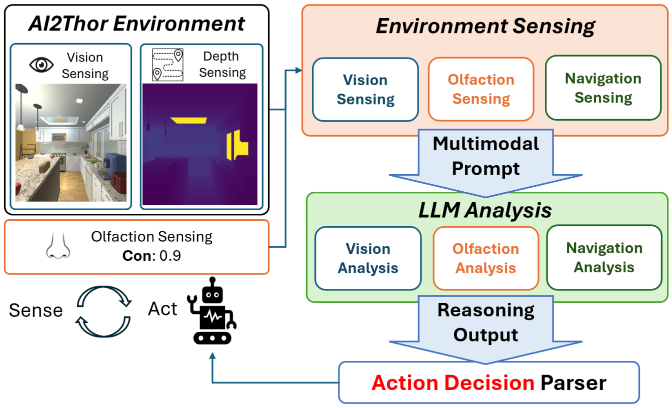

Fig. 2. The framework diagram of the proposed semantic OSL algorithm. The agent senses the environment and analyzes the vision, olfaction and navigation data to make decisions.

Semantic Analysis Instructions

Sequentially perform the following input analysis to select the best action:

Visual Analysis: do you see an object that is a possible source for {goal}?
- If yes, list the most likely object's name.
- If yes, and the front is obstacle free, move forward.

Olfactory Analysis:
- If the sensed concentration decreases, turn back.
- If the sensed concentration increases and front is obstacle free, move forward.
- If the sensed concentration stays the same, focus on visual and navigation analysis.

Navigation Analysis:
- If the front is blocked, turn to the side that is obstacle free and is more likely to lead you to the possible {goal} related object.
- If the last action was to turn (Action_id = 2 or 3), and the forward direction is blocked, then Turn Back.

Fig. 3. The semantic analysis instruction consists of vision analysis, olfaction analysis and navigation analysis stages.

toward the source location. Our goal is to find a function $\mathbf{F}$, which can produce effective robot actions $\mathbf{a}^t$ based on sensory observations $\mathbf{o}^t$. Mathematically, we can present our goal as:

$$\mathbf{a}^t = \mathbf{F}(\mathbf{o}^t). \tag{1}$$

In this work, robot actions $\mathbf{a}^t = \{a_1, a_2, a_3, a_4\}$ represent four discrete actions, i.e., Forward, Turn Left, Turn Right, and Turn Back. For observations $\mathbf{o}^t$, the visual observation v is captured from the robot's front-facing camera, and it is provided to LLM directly. The olfactory observation ρ is the odor concentration at the robot location, measured by the onboard chemical sensor. This information is provided to the LLM in a textual format.

B. Odor Distribution Model

In this work, we employed a Gaussian plume model [34] to simulate the odor plume distribution. To calculate the odor concentration ρ at the robot location $p = (x, y)$, we use the following equation:

$$\rho = \frac{q_s}{4\pi D\,|p - p_s|} \exp\left(\frac{-\Delta y \cdot U}{2D} + \frac{-|p - p_s|}{\lambda}\Delta t\right) \tag{2}$$

where

$$\begin{aligned} \lambda &= \sqrt{\frac{D\tau}{1 + \frac{U^2\tau}{4D}}} \\ \Delta y &= -(x - x_s)\sin(\psi) + (y - y_s)\cos(\psi). \end{aligned} \tag{3}$$

In the above model, q_s is the odor-releasing strength; D is the isotropic diffusivity; p_s is the odor source location; ψ is the wind direction; and τ is the odor particle lifetime. We only consider a single odor source, and since the search area is indoor, we assume zero external wind (i.e., U and ψ are zero) within the search area to simplify the problem.

C. LLM-based Semantic OSL Navigation

Fig. 2 illustrates the proposed semantic OSL framework, which includes three main blocks, i.e., Environmental Sensing, LLM Analysis, and Action Decoder.

1) Environmental Sensing: In the environmental sensing block, the robot gathers observations to describe its surroundings, including an RGB image of its front view, the odor concentration at its current position, and a depth image of its front view. Specifically, the RGB image is directly provided to the LLM, while the current and previous odor concentrations are converted into textual descriptions and fed into the LLM. The depth image is used to detect obstacles; if an object is detected within 0.7 m (the threshold defined in this work) of the robot, the LLM is informed that an obstacle is present in the front.

This environmental information is then used to construct the Multimodal Prompt, which comprises the collected observations, a system prompt, and analysis instructions. The system prompt outlines the robotic OSL task, defines the available actions, and provides relevant context. The analysis instructions guide the LLM to generate responses in a structured format using Step-by-Step Analysis [14]. The detailed LLM analysis instructions are presented as follows.

2) LLM Analysis: Fig. 3 shows the step-by-step reasoning instructions provided to the LLM. The analysis flow mirrors human odor search behavior. When an odor of interest is detected, visual processing is used to directly identify potential odor sources. If visual reasoning fails to pinpoint the object, olfactory cues guide the robot toward the odor's direction. The actions are chosen to avoid collision with surrounding obstacles in the environment.

In the **Vision Analysis**, the goal is to use visual reasoning to understand the surrounding environment of the agent and to visually pinpoint the potential odor source. To achieve this, the LLM is first tasked with detecting the overall scene (for example, identifying a kitchen) and recognizing the objects present (such as an oven) in the egocentric visual frame. The LLM model is then tasked to compare the semantic definition of the environment to the semantic definition of the sensed odor ({goal} in the prompt). The LLM model assesses whether any of these objects might be the source of the target odor. If a potential odor source is detected within the visual frame,

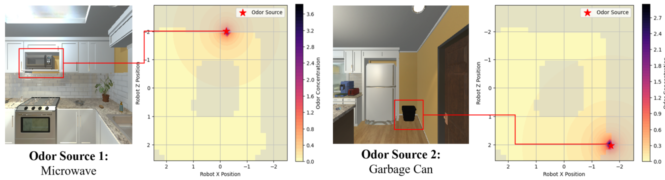

Fig. 4. The two odor source locations within the search area.

the LLM selects an action to move toward it.

In the **Olfactory Analysis**, the LLM is instructed to compare the current odor concentration value with the previously sensed concentration value. If the current odor concentration is greater than the previous time step, then it indicates that the robot is approaching the odor source, and it is instructed to continue this trajectory. On the other hand, decreasing concentration indicates the robot is moving away from the odor source, and thus it should reverse its trajectory.

In the **Navigation Analysis**, the LLM evaluates the surrounding obstacles, i.e., the information derived from depth sensing. The LLM model is also instructed to consider the consequence of the last selected action in planning the next action. This evaluation ensures that the chosen actions will allow the robot to approach the odor source without collisions.

3) Action Decoder: After the LLM's response, the Action Decoder will convert the textual LLM's response into an action ID, i.e., 1 to 4, where 1-Forward, 2-Turn Right, 3-Turn Left, and 4-Turn Back. Every time step, LLM is instructed to select only one action ID. Then, the robot will follow the selected action to update its position or orientation. It should be mentioned that Only the Forward action will change the robot's location, and the other actions will only change the robot's orientation.

IV. Experiment and Results

A. Experiment Setup

To evaluate the performance of the proposed OSL navigation algorithm, we implemented it in a simulated household environment using the AI2THOR [35] simulator. The simulated search environment is 5 m $\times$ 5 m. In experiments, we created two search scenarios, where the odor source is located at either the Microwave or the Garbage Can. The odor plume concentration at any location within the search area was calculated using Equation 2. For the Microwave, the odor description is "smoke", while for the Garbage Can, the odor description is "rotten smell".

Fig. 4 provides a top-down view of the odor plume distributions for both scenarios. We use the same parameters presented in [36] for our odor simulator, i.e., $q_s = 2000$ mg/s, $D = 10$ m^2/s, $\tau = 1000$ s, $\delta_t = 1$. During the search process, the robot is directed by the LLM to perform the generated action. A successful search is defined as the robot reaching the odor source location (i.e., the distance between the robot and the odor source is less than a threshold, 0.8 m in this work) within the decision-making step limit (i.e., 80 steps in this work). A failure occurs if the robot is unable to reach the odor source within this limit. In our experiments, the LLM model is GPT-4o.

B. Search Results

1) Search Trajectories: Fig. 5 illustrates the search trajectories of the proposed OSL navigation algorithm. In our implementation, we incorporated an adaptive speed control mechanism, where the robot adjusts its speed based on the sensed odor concentration: when the detected concentration is high, the robot slows down; when the concentration is low, the robot moves faster. The rationale behind this design is that when the robot is far from the odor source (i.e., low odor concentration), it should move quickly to explore the environment efficiently. Conversely, when it is near the odor source (i.e., high odor concentration), it should slow down to precisely locate the source. As shown in Fig. 5, the proposed semantic OSL algorithm enables the robot to successfully identify the odor source across different initial robot and odor source locations.

C. LLM Reasoning

Fig. 6 presents the reasoning results of the LLM based on the provided visual and olfactory inputs. The scenes, labeled S1–S4, correspond to different moments during the search process illustrated in Fig. 5. In S1, based on the detected odor type "Smoke", the LLM infers that the potential odor source is a "Stove" or "Oven". The directional inference is accurate, as the actual odor source, "Microwave," is located above the stove. Regarding the olfactory analysis, the robot detects an increase in odor concentrations, leading the LLM to conclude that it is moving toward the correct odor source. In the navigation analysis, the LLM is informed that there is no obstacle ahead and, given the visual and olfactory analysis, suggests Forward as the optimal action.

In S2, the robot does not detect any visually relevant objects associated with the odor source. However, an increase in sensed odor concentration suggests that the robot is moving in the right direction. As a result, the LLM advises continuing forward to approach the source. In S3, no useful visual objects are detected, and olfactory observations indicate a decrease in odor concentration, suggesting that the robot is moving away from the source. In response, the LLM determines that the best action is to turn back. In S4, neither useful visual nor olfactory observations are available, and the robot's front path is blocked. Given this scenario, the LLM directs the robot to turn right to continue exploring the search area.

D. Comparative Analysis

We compare the search performance of the proposed OSL navigation algorithm with three other methods: Random Walk, Vision-Only (VO), and Olfaction-Only (OO). In the Random Walk approach, the robot moves forward until it encounters an obstacle. Then, it randomly decides to turn left or right and continues forwarding. The VO and OO algorithms also employ the LLM for decision-making to generate a robot action command. However, in VO, the LLM does not receive

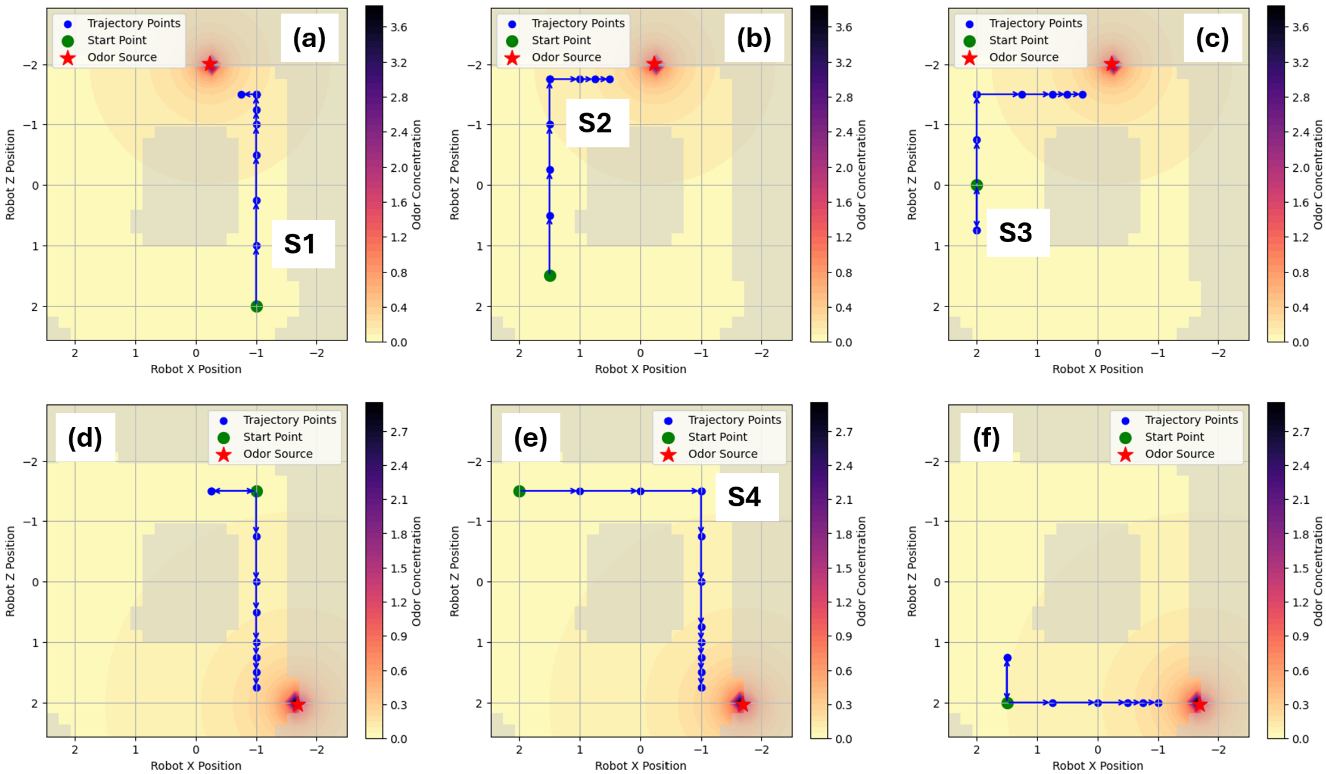

Fig. 5. Search trajectories of the proposed semantic OSL navigation algorithm. Plots (a)-(c) are trajectories with the odor source at "Microwave", and the detected odor type is "Smoke". (d)-(f) are trajectories with the odor source at "Garbage Can", and the detected odor type is "Rotten Smell". S1-S4, labeled alongside the trajectories are sample scenes, where the agent's decision-making results are presented in Fig. 6. The robot's moving direction is labeled by the arrows. In all searches, the robot successfully reached the odor source location.

TABLE I
THE COMPARATIVE ANALYSIS RESULTS OF THE PROPOSED OSL NAVIGATION ALGORITHM AND OTHER ALGORITHMS

	Avg. Steps ↓	Avg. Dist. ↓	Success Rate ↑
Random Walk	58.2	22.8 m	33%
Vision Only (VO)	47.7	9.6 m	83.3%
Olfaction Only (OO)	35.7	10.3 m	100%
The Proposed Method	**11.3**	**4.6 m**	**100%**

odor concentrations, while in OO, visual information (i.e., the image captured from the front-facing camera) is not provided. All robot initial locations and odor source locations remain the same as those presented in Fig. 5.

Table I presents the search performance results of the different OSL navigation algorithms. The proposed semantic OSL navigation algorithm outperforms all other methods across evaluation metrics, including average decision-making steps, average travel distance, and success rate (out of six trials per method). Using Random Walk as a baseline, we observe that LLM-based decision-making significantly improves search performance. However, when only one sensory modality is used, as in VO and OO, the robot's search efficiency declines compared to the integrated method.

Specifically, the VO approach struggles in cases where the robot lacks meaningful visual cues to guide the search (e.g., S2 in Fig. 6), leading to a lower success rate. The OO approach achieves a 100% success rate, but without visual information, the robot takes longer to locate the source, resulting in suboptimal search performance. The proposed method, integrating both visual and olfactory observations, achieves a 100% success rate and the shortest traveling distance, compared to other methods. These results highlight the advantage of integrating semantic visual and olfactory information, demonstrating that a multi-modal approach improves both the effectiveness and efficiency of robotic OSL tasks.

The primary objective of our work is to demonstrate semantic Odor Source Localization (OSL) navigation by integrating both olfactory and visual sensory modalities. Many existing OSL algorithms (such as those based on RL or DL) primarily focus on olfactory-only methods. Our experiments show that integration of vision with olfaction outperforms single sensory modality-based vision-only and olfactory-only methods.

V. CONCLUSION AND FUTURE WORKS

This paper presents a semantic OSL navigation algorithm that integrates both visual and olfactory observations to enhance OSL tasks. We leverage an LLM as the decision-making agent to analyze sensory inputs, infer the possible odor source location, and guide the robot toward it. Additionally, we

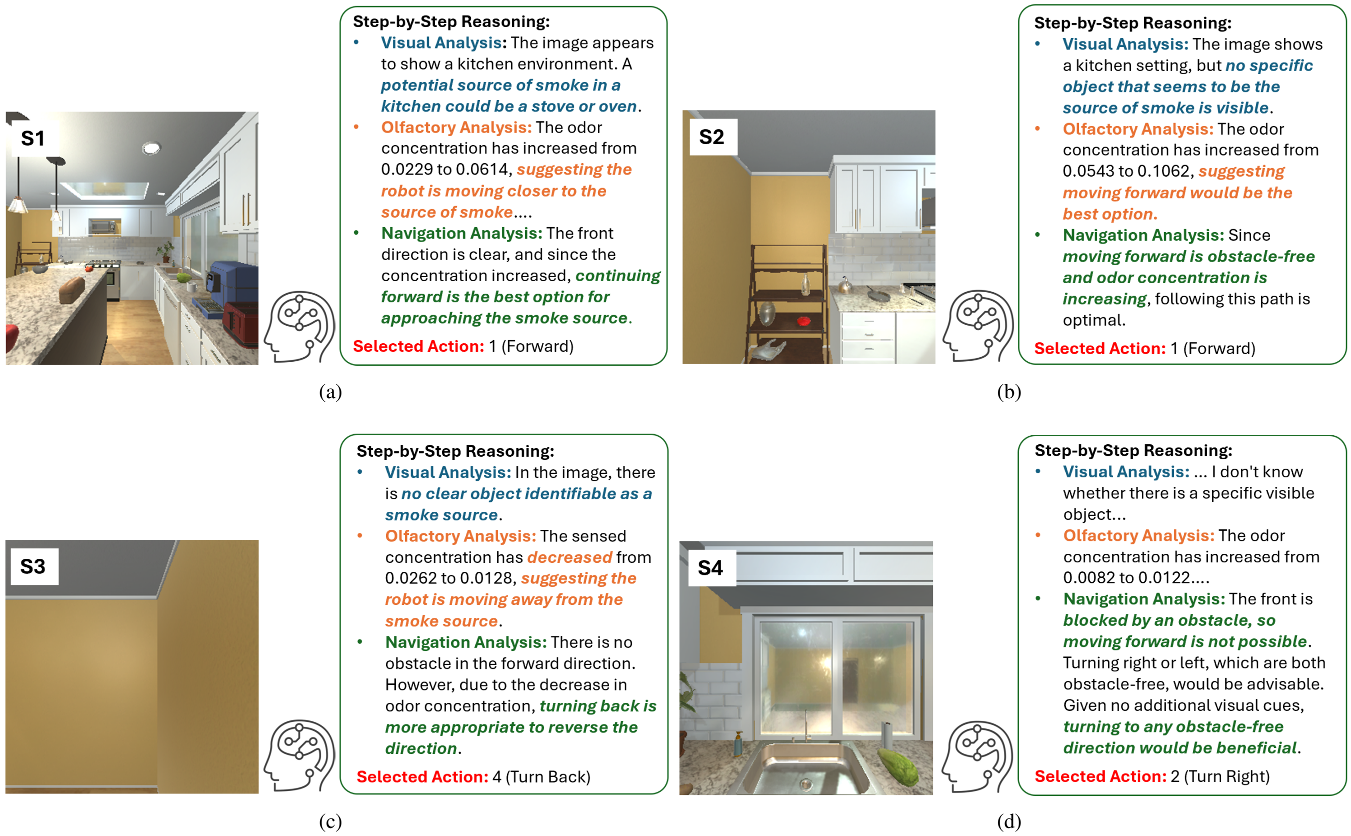

Fig. 6. The reasoning outputs of the LLM with different visual and olfactory inputs. The scenes for obtaining S1-S4 are labeled in Fig. 5. For S1 and S2, the odor source is the "Microwave", and for S3 and S4, the odor source is "Garbage Can". Key reasoning analysis for **Visual Analysis**, **Olfactory Analysis**, and **Navigation Analysis** are labeled in Dark Teal, Orange, and Dark Green, respectively.

employ Chain-of-Thought reasoning to help the LLM exploit semantic associations between detected odors and observed visual targets. The proposed OSL navigation algorithm is implemented in a simulated household environment, where we evaluate its performance across two odor source search scenarios and three robot initial positions, comparing it against Random Walk, Vision-Only (VO), and Olfaction-Only (OO) algorithms. Simulation results demonstrate that the proposed multi-modal OSL navigation algorithm achieves the highest success rate and shortest travel distance among all tested methods.

The main innovation of this project lies in integrating vision and olfaction through a multimodal LLM. The multimodal GPT-4 model is capable of reasoning about the robot's next actions based on current visual and olfactory observations, using common-sense knowledge. Our experiments indicate that highly sophisticated reasoning is not required to guide the robot effectively toward the odor source. Consequently, the multimodal LLM was able to infer potential odor sources in the environment, select appropriate exploratory actions, and ultimately localize the true odor source with reliable performance.

In future works, we plan to further explore semantic OSL and the integration of visual and olfactory observations in two directions. First, we will extend the approach to multi-source searching scenarios, where multiple odor sources exist in the environment and emit odor plumes simultaneously. This extension will enable the navigation algorithm to operate in more complex search environments with broader real-world applications. Second, we will implement the proposed navigation algorithm into real-world tests. While the simulator used in this study provides a highly realistic environment, a gap remains between simulation and real-world conditions. Future work will focus on implementing the proposed navigation algorithm on a physical robot to evaluate its performance in real-world search environments.

References

[1] X.-x. Chen and J. Huang, "Odor source localization algorithms on mobile robots: A review and future outlook," *Robotics and Autonomous Systems*, vol. 112, pp. 123–136, Feb. 2019.

[2] T. Jing, Q.-H. Meng, and H. Ishida, "Recent Progress and Trend of Robot Odor Source Localization," *IEEJ Transactions on Electrical and Electronic Engineering*, vol. 16, no. 7, pp. 938–953, Jul. 2021.

[3] J. Wang, Y. Lin, R. Liu, and J. Fu, "Odor source localization of multi-robots with swarm intelligence algorithms: A review," *Frontiers in Neurorobotics*, vol. 16, p. 949888, 2022.

[4] S. Shigaki, Y. Yoshimura, D. Kurabayashi, and K. Hosoda, "Palm-sized quadcopter for three-dimensional chemical plume tracking," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1–12, 2022.

[5] W. Li, J. A. Farrell, S. Pang, and R. M. Arrieta, "Moth-inspired chemical plume tracing on an autonomous underwater vehicle," *IEEE Transactions on Robotics*, vol. 22, no. 2, pp. 292–307, 2006.
[6] D. Rjoub, A. Alsharoa, and A. Masadeh, "Unmanned-aircraft-system-assisted early wildfire detection with air quality sensors," *Electronics*, vol. 12, no. 5, p. 1239, 2023.
[7] L. Wang, S. Pang, M. Noyela, K. Adkins, L. Sun, and M. El-Sayed, "Vision and olfactory-based wildfire monitoring with uncrewed aircraft systems," in *2023 20th International Conference on Ubiquitous Robots (UR)*. IEEE, 2023, pp. 716–723.
[8] J. Burgués and S. Marco, "Environmental chemical sensing using small drones: A review," *Science of The Total Environment*, vol. 748, p. 141172, Dec. 2020.
[9] J. Thoma, D. P. Paudel, A. Chhatkuli, T. Probst, and L. V. Gool, "Mapping, localization and path planning for image-based navigation using visual features and map," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 7383–7391.
[10] S. Kuang and T. Zhang, "Smelling directions: olfaction modulates ambiguous visual motion perception," *Scientific reports*, vol. 4, no. 1, pp. 1–5, 2014.
[11] T. Brown, B. Mann, N. Ryder, M. Subbiah, J. D. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell *et al.*, "Language models are few-shot learners," *Advances in neural information processing systems*, vol. 33, pp. 1877–1901, 2020.
[12] A. Radford, J. W. Kim, C. Hallacy, A. Ramesh, G. Goh, S. Agarwal, G. Sastry, A. Askell, P. Mishkin, J. Clark *et al.*, "Learning transferable visual models from natural language supervision," in *International conference on machine learning*. PMLR, 2021, pp. 8748–8763.
[13] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: unified, real-time object detection (2015)," *arXiv preprint arXiv:1506.02640*, vol. 825, 2015.
[14] J. Wei, X. Wang, D. Schuurmans, M. Bosma, F. Xia, E. Chi, Q. V. Le, D. Zhou *et al.*, "Chain-of-thought prompting elicits reasoning in large language models," *Advances in neural information processing systems*, vol. 35, pp. 24 824–24 837, 2022.
[15] H. Ishida, K.-i. Suetsugu, T. Nakamoto, and T. Moriizumi, "Study of autonomous mobile sensing system for localization of odor source using gas sensors and anemometric sensors," *Sensors and Actuators A: Physical*, vol. 45, no. 2, pp. 153–157, 1994.
[16] H. Ishida, G. Nakayama, T. Nakamoto, and T. Moriizumi, "Controlling a gas/odor plume-tracking robot based on transient responses of gas sensors," *IEEE Sensors Journal*, vol. 5, no. 3, pp. 537–545, 2005.
[17] A. Lilienthal and T. Duckett, "Experimental analysis of gas-sensitive Braitenberg vehicles," *Advanced Robotics*, vol. 18, no. 8, pp. 817–834, 2004.
[18] S. Shigaki, S. Haigo, C. H. Reyes, T. Sakurai, R. Kanzaki, D. Kurabayashi, and H. Sezutsu, "Analysis of the role of wind information for efficient chemical plume tracing based on optogenetic silkworm moth behavior," *Bioinspiration & biomimetics*, vol. 14, no. 4, p. 046006, 2019.
[19] A. M. Matheson, A. J. Lanz, A. M. Medina, A. M. Licata, T. A. Currier, M. H. Syed, and K. I. Nagel, "A neural circuit for wind-guided olfactory navigation," *Nature Communications*, vol. 13, no. 1, p. 4613, 2022.
[20] B. T. Michaelis, K. W. Leathers, Y. V. Bobkov, B. W. Ache, J. C. Principe, R. Baharloo, I. M. Park, and M. A. Reidenbach, "Odor tracking in aquatic organisms: the importance of temporal and spatial intermittency of the turbulent plume," *Scientific reports*, vol. 10, no. 1, p. 7961, 2020.
[21] L. Wang, S. Pang, and J. Li, "Olfactory-based navigation via model-based reinforcement learning and fuzzy inference methods," *IEEE Transactions on Fuzzy Systems*, vol. 29, no. 10, pp. 3014–3027, 2021.
[22] S. Pang and J. A. Farrell, "Chemical plume source localization," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 36, no. 5, pp. 1068–1080, 2006.
[23] J. Li, Q. 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.
[24] H. Jiu, Y. Chen, W. Deng, and S. Pang, "Underwater chemical plume tracing based on partially observable markov decision process," *International Journal of Advanced Robotic Systems*, vol. 16, no. 2, p. 1729881419831874, 2019.
[25] S. Pang and F. Zhu, "Reactive planning for olfactory-based mobile robots," in *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2009, pp. 4375–4380.
[26] L. Wang and S. Pang, "Chemical plume tracing using an auv based on pomdp source mapping and a-star path planning," in *OCEANS 2019 MTS/IEEE SEATTLE*. IEEE, 2019, pp. 1–7.
[27] J. Wang, Z. Wu, Y. Li, H. Jiang, P. Shu, E. Shi, H. Hu, C. Ma, Y. Liu, X. Wang *et al.*, "Large language models for robotics: Opportunities, challenges, and perspectives," *arXiv preprint arXiv:2401.04334*, 2024.
[28] S. Y. Gadre, M. Wortsman, G. Ilharco, L. Schmidt, and S. Song, "Cows on pasture: Baselines and benchmarks for language-driven zero-shot object navigation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 23 171–23 181.
[29] N. Yokoyama, S. Ha, D. Batra, J. Wang, and B. Bucher, "Vlfm: Vision-language frontier maps for zero-shot semantic navigation," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2024, pp. 42–48.
[30] V. S. Dorbala, G. Sigurdsson, R. Piramuthu, J. Thomason, and G. S. Sukhatme, "Clip-nav: Using clip for zero-shot vision-and-language navigation," *arXiv preprint arXiv:2211.16649*, 2022.
[31] R. Schumann, W. Zhu, W. Feng, T.-J. Fu, S. Riezler, and W. Y. Wang, "Velma: Verbalization embodiment of llm agents for vision and language navigation in street view," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 38, no. 17, 2024, pp. 18 924–18 933.
[32] D. Shah, B. Osinski, B. Ichter, and S. Levine, "Robotic navigation with large pre-trained models of language," *Vision, and Action*, 2022.
[33] P. Chen, X. Sun, H. Zhi, R. Zeng, T. H. Li, G. Liu, M. Tan, and C. Gan, "A^2 nav: Action-aware zero-shot robot navigation by exploiting vision-and-language ability of foundation models," *arXiv preprint arXiv:2308.07997*, 2023.
[34] M. Vergassola, E. Villermaux, and B. I. Shraiman, "'infotaxis' as a strategy for searching without gradients," *Nature*, vol. 445, no. 7126, pp. 406–409, 2007.
[35] E. Kolve, R. Mottaghi, W. Han, E. VanderBilt, L. Weihs, A. Herrasti, M. Deitke, K. Ehsani, D. Gordon, Y. Zhu *et al.*, "Ai2-thor: An interactive 3d environment for visual ai," *arXiv preprint arXiv:1712.05474*, 2017.
[36] M. Park, P. Ladosz, and H. Oh, "Source term estimation using deep reinforcement learning with gaussian mixture model feature extraction for mobile sensors," *IEEE Robotics and Automation Letters*, vol. 7, no. 3, pp. 8323–8330, 2022.