An Intuitive Voice Interface for a UV Disinfecting Mobile Manipulator Robot

Alan G. Sanchez, William D. Smart
Collaborative Robotics and Intelligent Systems (CoRIS) Institute, Oregon State University
{sanchala, bill.smart}@oregonstate.edu

Abstract

In the wake of the global health crises caused by the COVID-19 pandemic, there is a pressing need for innovative disinfection methods that are both effective and user-friendly to a broad user base. This paper introduces an approach that allows a user to instruct tasks to a UV disinfection robot via speech. The implementation of a voice interface offers a hands-free operation and caters to non-technical users who require a simple and effective way to command the robot. Through a combination of object recognition, natural language processing using a large language model (LLM), and task planning, our system has the potential to execute tasks more effectively since it is more context-aware of its sanitizing duties.

1. Introduction

Since the COVID-19 pandemic, research institutions worldwide have focused on expanding the capabilities of UV disinfection robots. These efforts have been directed towards several key areas, including design enhancements[4, 9, 14, 16], optimization of UV coverage [10, 15], and measuring the UV dose accumulation across a surface [5, 7, 12]. However, while these technological advances in UV disinfection robotics have been substantial, they primarily address UV disinfection's mechanical and measurement aspects. The challenge remains in enabling these robots to understand the context of the spaces and objects they are tasked with disinfecting. This is where the potential of embodied AI can help. Leveraging the capabilities of speech-to-text interfaces and large language models (LLM), robots can process and interpret human voice commands, making them more context-aware.

Our current UV disinfection robot system has multiple operational modes. It can follow a trajectory generated by our path planner, or an operator can guide the manipulator directly, where the human-guided path is recorded and replayed back, also known as Program by Demonstration (PbD). The PbD feature enables humans to guide the robot to disinfect items with semantically informed paths, potentially improving coverage [11]. In this paper, we describe our plans to augment our interface by including the following features: (1) Detect objects on a table or within a targeted area, (2) Receive and interpret natural-language commands, (3) Formulate a plan to disinfect items in sequence, and (4) Request assistance for any unfamiliar tasks. These functionalities result from the recent development in language processing and object detection technologies. Adding these features can improve our robot's capabilities to make it a practical and user-friendly system.

2. Related

In recent years, research has increasingly integrated AI with mobile manipulators to enable robots to understand and execute tasks independently. As a result, this has led to the development of frameworks that improve Open-Vocabulary robots and their capabilities to interact with human-centered environments[2, 3, 8, 13]. Liu et al. [6] introduces an open-knowledge-based framework, *OK-Robot*, that integrates Vision-Language Models, navigation primitives, and grasping primitives to perform pick-and-place operations in real-world environments while not requiring training. The study evaluated OK-Robot by conducting 171 pick-and-place experiments in 10 different environments. The framework completed full pick-and-place in 58 percent of cases in unseen and untrained environments.

Collins et al. [1] showcases a system, ForceSight, that predicts visual-force goals from input images and text prompts. ForceSight key contributions are its ability to process an RGBD image and a text prompt to produce an affordance map, suitable end-effector locations, and a force goal with target grip and applied forces. The framework was evaluated on a Stretch Mobile Manipulator robot for ten mobile manipulation tasks and had a success rate of 81 percent over 100 pick-and-place, drawer opening, and light switch flipping tasks. Future work explores how a robot can execute more complex tasks by breaking them into subtasks. In this paper, we lay out our plans for developing a voice-command interface for a UV disinfection robot. In the following section, we delve into the design layout of our project.

3. Design Layout

To begin a disinfection task, a human supervisor provides a verbal command for a Fetch mobile manipulator robot to execute. For example, if a human supervisor says, "Fetch, please disinfect the eating utensils on the table," this verbal command is processed and initiates other functionalities.

When a verbal command is given, the robot employs its head camera to detect objects placed on the table and record their arrangement. The visual information gathered by the camera, combined with the context of the human-supplied verbal instruction, forms the inputs that are added to the prompt, a set of instructions to an LLM. This prompt encapsulates both the specifics of the task at hand (disinfecting utensils, in this case) and the spatial arrangement of these utensils that the robot perceives.

Next, the prompt is submitted to the LLM, and the robot processes the request and produces a structured representation of the disinfection task. This representation includes a sequence of objects to disinfect, including their corresponding location. A traditional task executive is then utilized to produce a tool path to carry out the disinfection process.

Further, the robot is responsive to the user if there are any operational constraints. If it encounters an object for which it does not have a predefined disinfection trajectory, the robot will notify the human supervisor for additional help. It does this by communicating to the human to guide its arm around the object of interest; simultaneously, the robot will record the human-guided path and store it in the repository of other trajectories. This feature ensures the robot can handle a wide range of objects and scenarios, even those for which it has not been explicitly trained.

4. Human Centered Study

We plan to incorporate a human-centered study, allowing us to observe and analyze how users interact with the robot during disinfection tasks. The primary goal is to focus on the human's decisions and actions when guiding the robot's arm to disinfect objects, providing insights into user priorities and high-level reasoning.

Our study will have a variety of setups featuring a collection of objects and layout arrangements. We aim to focus on several metrics, which include the sequence of objects chosen for disinfection, the reason behind a participant's decision-making, and observe how users guide the robot's arm when moving around objects. By examining users' patterns, we can recognize the prioritized objects based on their perceived importance, ease of disinfection, or potential contamination risk. We can also identify if the physical characteristics of objects (e.g., size, height, material) influence how and why they are disinfected. Further, noting any difficulties or hesitations could provide areas for improvement when developing the repository of joint trajectories. The study's findings can also improve the analysis and the robot's design layout. For example, if the study reveals that users consistently prefer to disinfect certain types of objects first, the robot's programming could be adjusted to prioritize these objects in future disinfection tasks. From this analysis, we will extract any correlations between object characteristics and disinfection decisions and identify potential areas for enhancing the robot's context awareness and usability.

4.1. Evaluation

A metric to consider is the robot's success rate in executing disinfection tasks. We define success as the robot's ability to clean designated objects accurately based on verbal commands. This also means that if the robot either disinfects an additional item or neglects it, it is considered a failed attempt. Calculating the ratio of successfully completed tasks gauges the system's effectiveness. The success rate also reflects the robot's ability to understand the instructions a human supervisor provides since it produces what sequence of actions the robot must execute. Thus, evaluating the LLM's ability to interpret voice commands into executable tasks correctly is central to the robot's performance. To assess this, we will examine the LLM's accuracy in parsing various phrases of commands to test its comprehension capabilities. This evaluation can help us understand if modifications to our prompt can refine and improve our system. This project is still in the planning phases, and our primary focus is to establish a comprehensive evaluation framework that accurately measures our system's performance and effectiveness.

5. Discussion

In this paper, we presented a voice-command interface that utilizes an LLM to help a UV disinfection robot understand the context of the spaces and objects it is tasked to sanitize. The real world presents variability, making it challenging for a robot to execute tasks for immutable pre-defined trajectories. This also presents a hurdle for users who may not have the time and technical expertise to operate a complex robot system. Thus, simplifying the interaction to something as natural as voice communication can enhance the disinfection processes without having additional operational burdens. Our approach can potentially make UV disinfection robots more effective and user-friendly while exploring new methods to integrate human-robot interaction to address public health challenges.

References

[1] Jeremy A Collins, Cody Houff, You Liang Tan, and Charles C Kemp. Forcesight: Text-guided mobile manipulation with visual-force goals. *arXiv preprint arXiv:2309.12312*, 2023. 1

[2] Sergio Guadarrama, Erik Rodner, Kate Saenko, Ning Zhang, Ryan Farrell, Jeff Donahue, and Trevor Darrell. Open-vocabulary Object Retrieval. In *Robotics: science and systems*, page 6, 2014. 1

[3] Sergio Guadarrama, Erik Rodner, Kate Saenko, and Trevor Darrell. Understanding object descriptions in robotics by open-vocabulary object retrieval and detection. *The International Journal of Robotics Research*, 35(1-3):265–280, 2016. 1

[4] Moez Guettari, Ines Gharbi, and Samir Hamza. UVC disinfection robot. *Environmental Science and Pollution Research*, 28:40394–40399, 2021. 1

[5] Irfan Tito Kurniawan and Widyawardana Adiprawita. A method of ultraviolet-C surface irradiation simulation and evaluation. In *2021 International Symposium on Electronics and Smart Devices (ISESD)*, pages 1–5. IEEE, 2021. 1

[6] Peiqi Liu, Yaswanth Orru, Chris Paxton, Nur Muhammad Mahi Shafiullah, and Lerrel Pinto. Ok-robot: What really matters in integrating open-knowledge models for robotics. *arXiv preprint arXiv:2401.12202*, 2024. 1

[7] Joao Marcos Correia Marques, Ramya Ramalingam, Zherong Pan, and Kris Hauser. Optimized coverage planning for UV surface disinfection. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pages 9731–9737. IEEE, 2021. 1

[8] Andrew Melnik, Michael Büttner, Leon Harz, Lyon Brown, Gora Chand Nandi, Arjun PS, Gaurav Kumar Yadav, Rahul Kala, and Robert Haschke. UniTeam: Open Vocabulary Mobile Manipulation Challenge. *arXiv preprint arXiv:2312.08611*, 2023. 1

[9] Alyssa Pierson, John W Romanishin, Hunter Hansen, Leonardo Zamora Yañez, and Daniela Rus. Designing and deploying a mobile UVC disinfection robot. In *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 6700–6707. IEEE, 2021. 1

[10] Thomas Rubaek, Merima Cikotic, and Simon Falden. Evaluation of the UV-disinfection robot. *Blue Ocean Robotics*, 2016. 1

[11] Alan G Sanchez, Nash Bernhart, and William D Smart. More effective robotic uv disinfection of objects through human guidance. In *Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*, pages 930–934, 2024. 1

[12] Alan G. Sanchez, Nash Bernhart, and William D. Smart. More effective robotic uv disinfection of objects through human guidance. In *Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*, pages 930–934, 2024. 1

[13] Nur Muhammad Mahi Shafiullah, Anant Rai, Haritheja Etukuru, Yiqian Liu, Ishan Misra, Soumith Chintala, and Lerrel Pinto. On bringing robots home. *arXiv preprint arXiv:2311.16098*, 2023. 1

[14] Avni Shah. Robotic Arms Extend the Reach of UV Disinfection, 2021. 1

[15] Luca Tiseni, Domenico Chiaradi, Massimiliano Gabardi, Massimiliano Solazzi, Daniele Leonardis, and Antonio Frisoli. UV-C Mobile Robots With Optimized Path Planning. *Robotics & Automation Magazine*, 28(1):59–70, 2021. 1

[16] A Vyshnavi, A Manasa, Ch Hamsika, and P Shalini. Uv disinfection robot with automatic switching on human detection. *Eai endorsed transactions on internet of things*, 6 (23), 2020. 1