



Extracting Human Visual Perspective From Robot Sensors



Albert Chen^{†1}, Helena Sieh^{†1}, Wonse Jo², Lionel Robert², Dawn Tilbury²

¹ College of Engineering, ² Robotics Department, University of Michigan, Ann Arbor, MI



†: contributed equally to this work

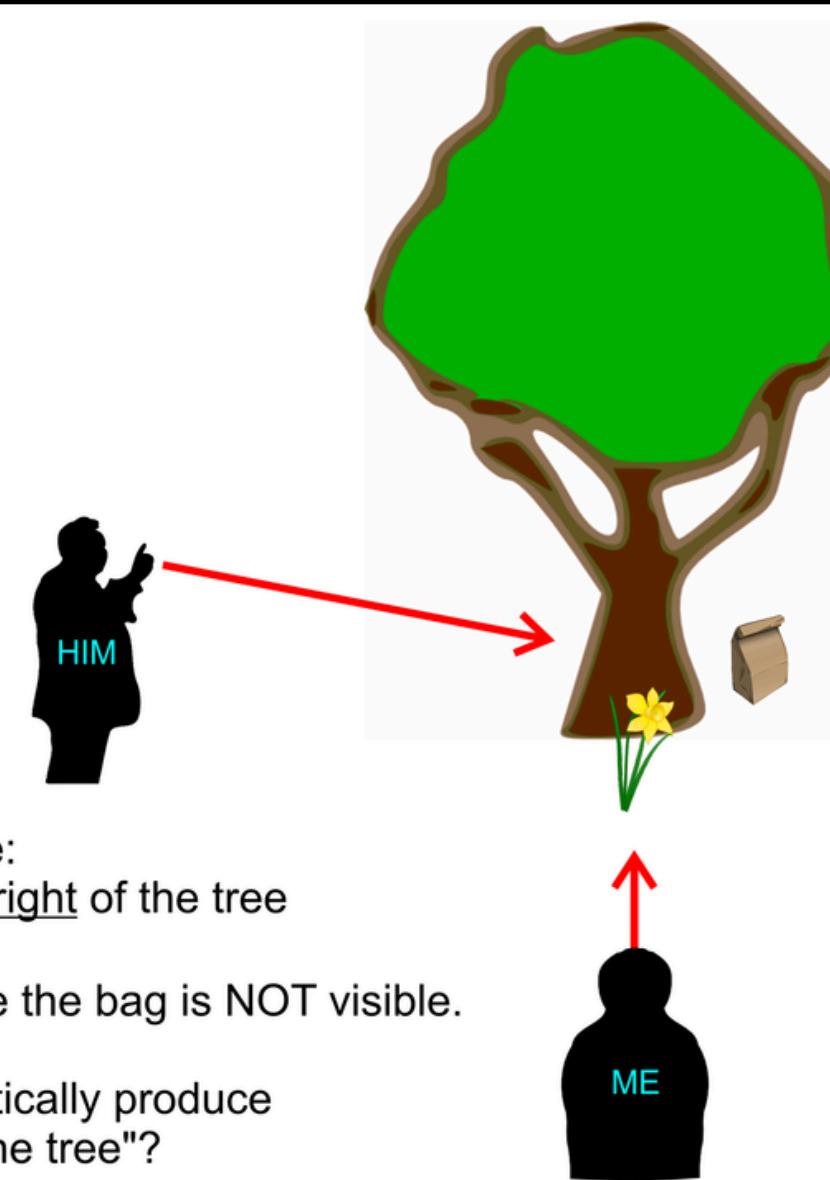
Introduction

Visual Perspective Taking (VPT): ability to understand another person's viewpoint, taking into account what they see and how they see it.

Why VPT?

- Contributes to better decision-making and social connections in HRI.
- Enables predictions with detailed and insightful observations.

Background



'PT-2: From his perspective:
The flower is on the right of the tree

'PT-1: From his perspective the bag is NOT visible.

Does VPT-1 automatically produce
"The bag is behind the tree"?

<Types of visual perspective tasking>

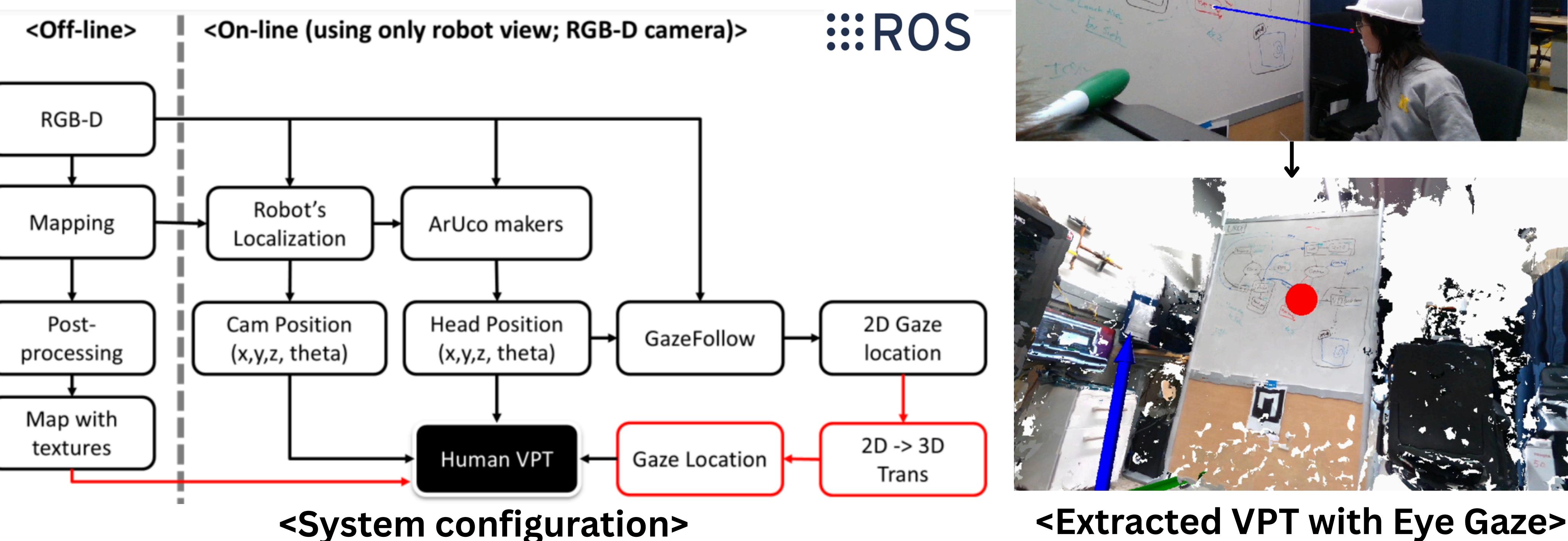
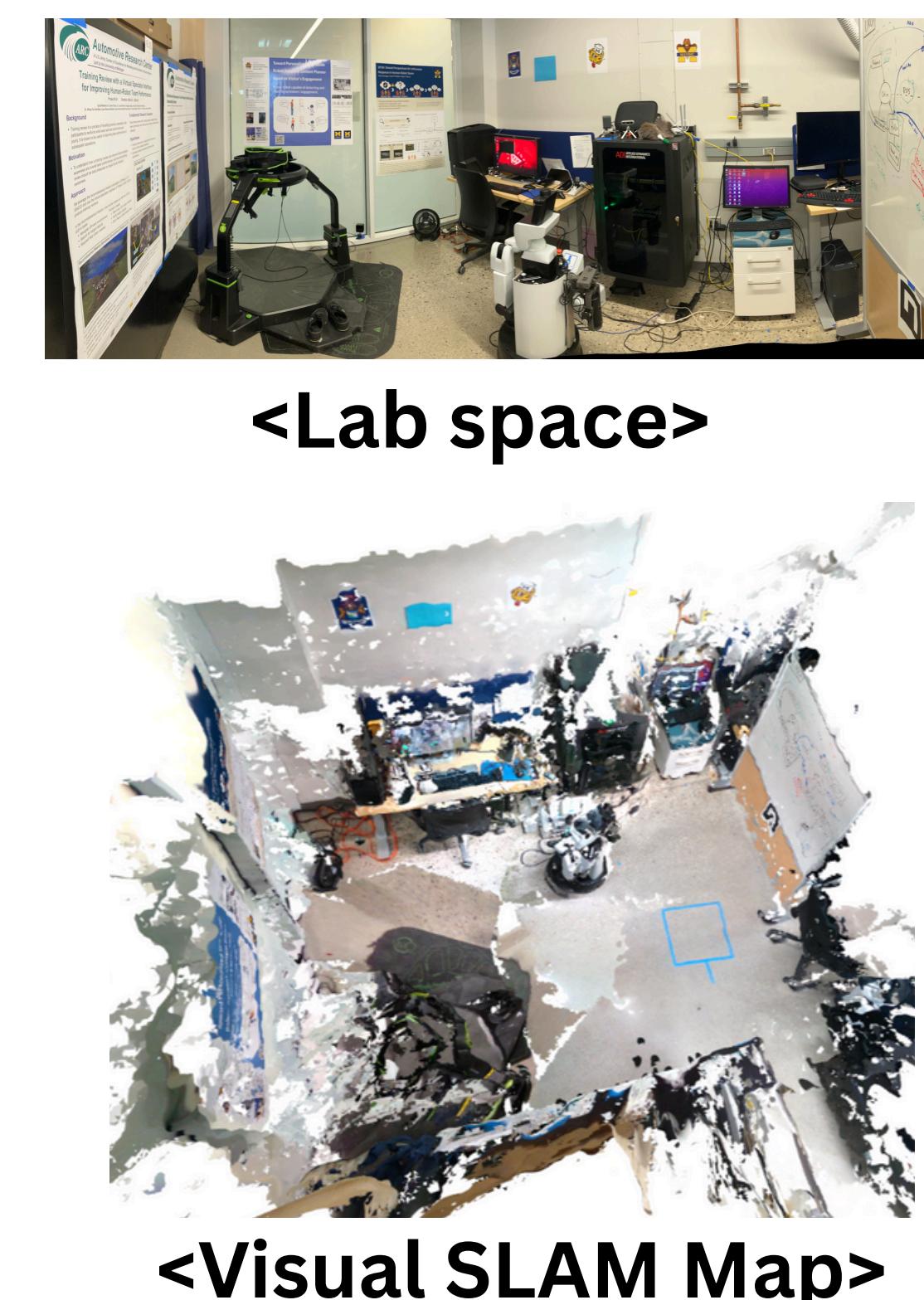
- Environment context allows us to understand which objects human can see and cannot see.
- But, there is less development in robot systems to understand human VPT.

Research Question

How can robots be designed to estimate human visual perspective taking (VPT) from common sensors?

Methods

- We propose a new multi modal human VPT extraction algorithm
- Uses Visual Simultaneous Localization and Mapping (VSLAM) to map out environment and RGB-D data to determine head and gaze positions.

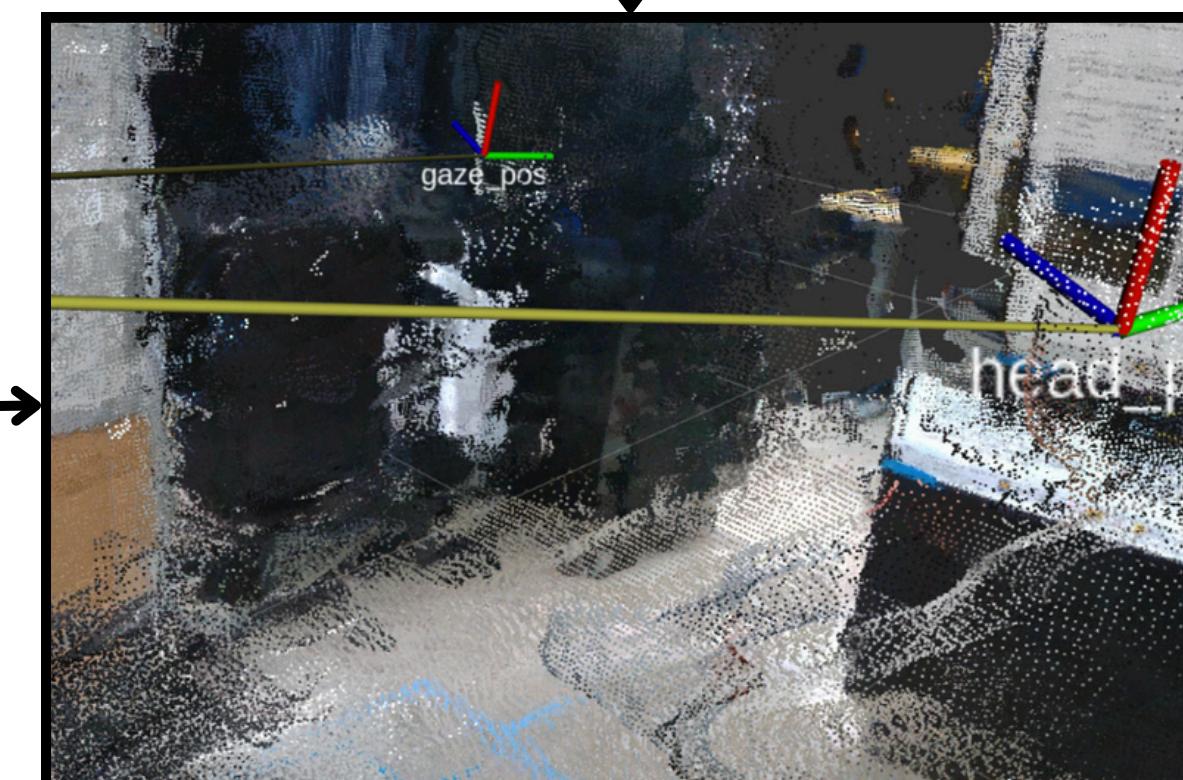
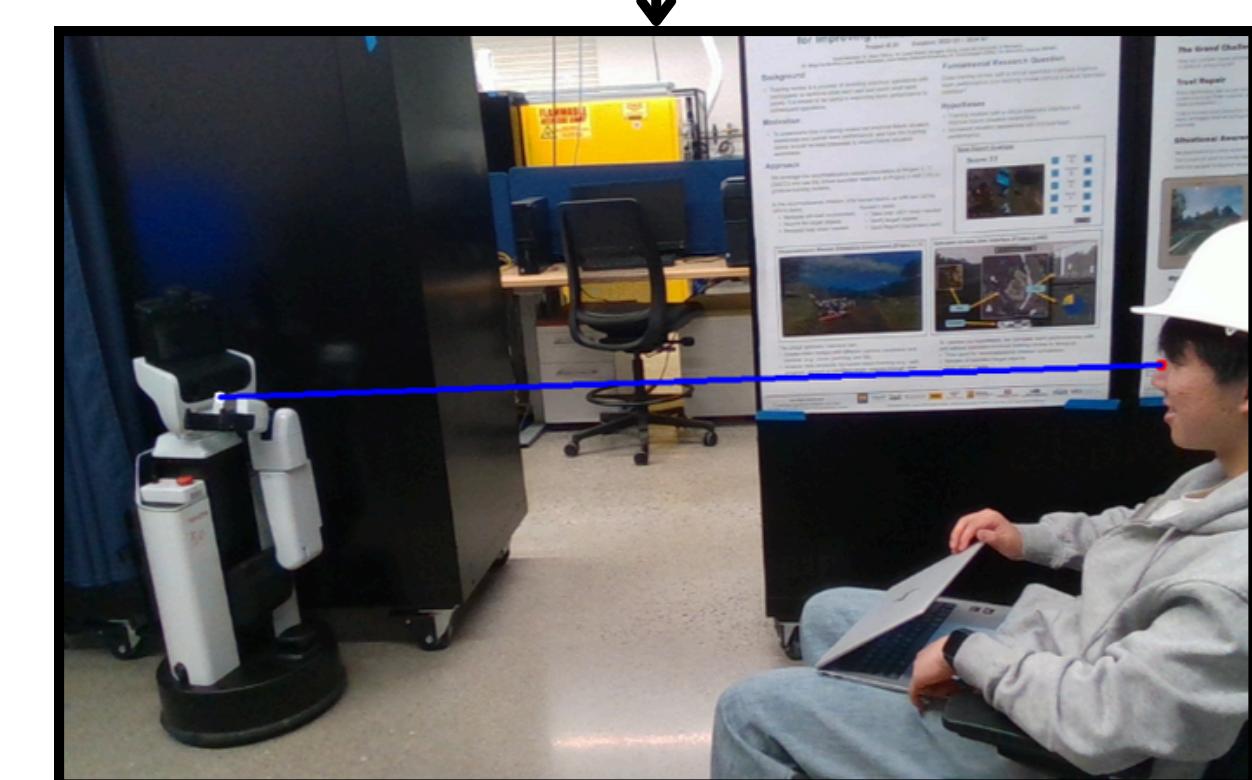
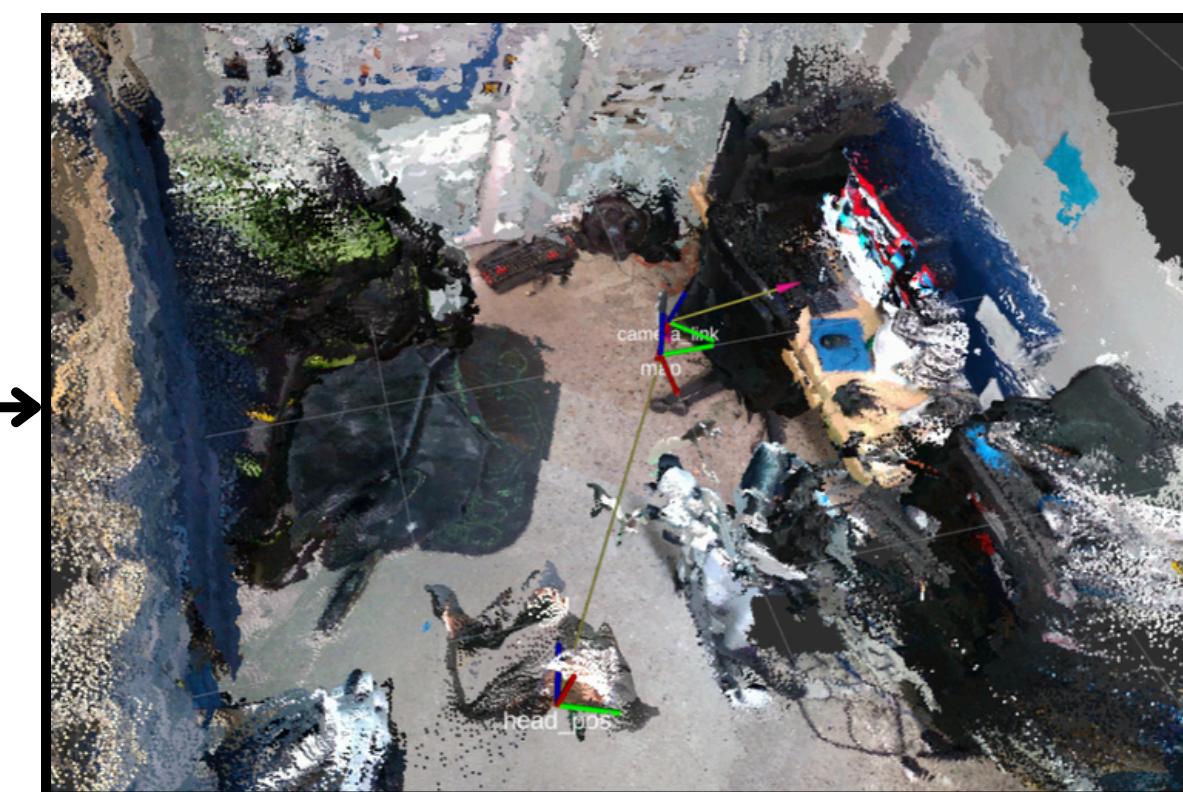


Results

Head Pose



Head Location on Map



Eye Gaze

Gaze Location on Map

- Estimates direction of the human's VPT with an image model and head position.
- Algorithm was validated through simulated and real-world testing

Conclusion & Future Works

- The algorithm offers a reliable method for robots to understand human perspectives in real-time.
- Future work will focus on improving the algorithm's real-world applications

More details can be found at the QR codes: Demo video, our VPT ROS package, and website.



<VPT Demo>



<Github>



<Website>

Acknowledgements

This work was funded by the Undergraduate Research Opportunity Program (UROP). Special thanks to Wonse Jo for mentorship and guidance throughout the project.