Leveraging Robot Gaze to Signal Membership in Dynamic Group Settings

Alina Shi, YSP Student, *Belmont High School*
Ilyas Mohamed, YSP Student, *Cambridge Rindge and Latin School*
Reina Chan, Computer Science, *Northeastern University*
Drake Moore, Computer Engineering, *Northeastern University*
Xiang Zhi Tan, Computer Science, *Northeastern University*

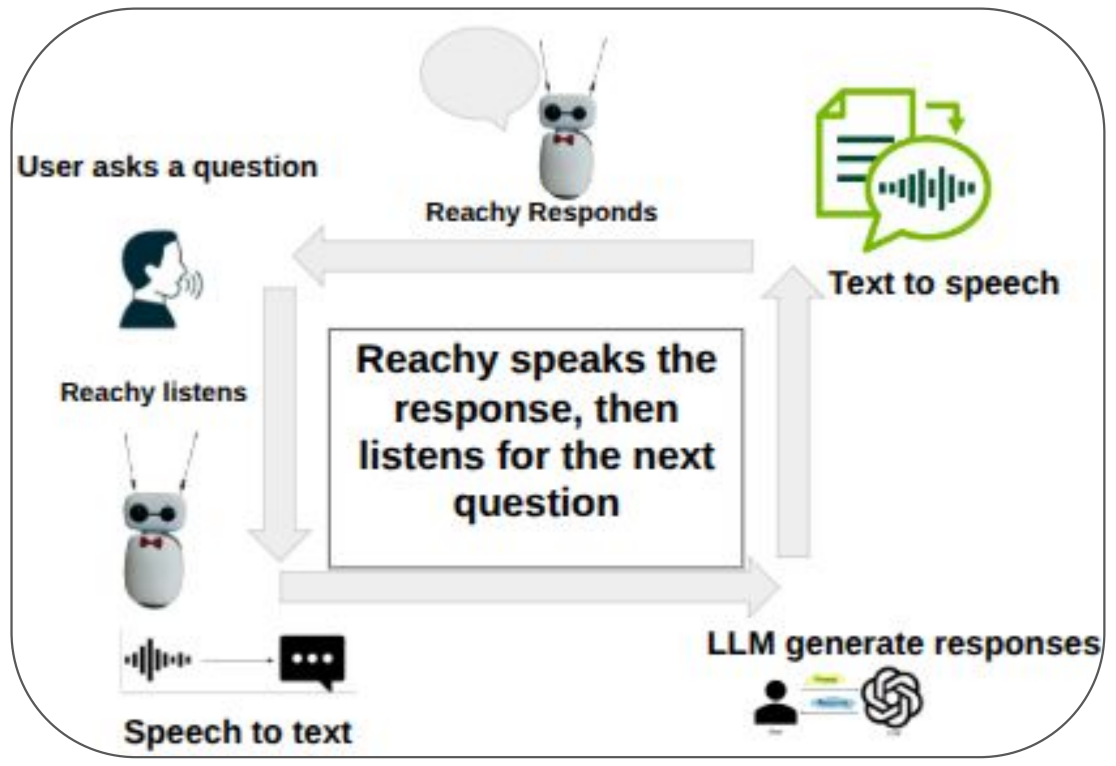
PARCS People And Robot Collaborative Systems Lab — Young Scholars Program — NSF — Northeastern University Khoury College of Computer Sciences — Northeastern University Michael B. Silevitch and Claire J. Duggan Center for STEM Education

Abstract

Humans use gaze to signal different manners such as thinking, paying attention to people and directing attention to other objects around them. Prior work has shown that humans naturally anthropomorphize robots and in order to maintain engagement, both humans and robots must communicate through gaze, dialogue, and movement. We investigate how these three components work together to shape interaction with groups of people, and how they can be implemented to improve the social interaction of a robot. The goal is to utilize the gaze of our robot to increase engagement with dynamic audiences.

Our project explores how a tabletop robot can use a local Large Language Model (LLM) with custom prompts and face tracking models to create more natural interaction behaviors. Once completed, we hope to deploy the robot in a building so it can interact with interested visitors and tour groups. In doing so, we will evaluate how different gaze behaviors can influence newcomer permanence and interaction in groups.

Experimental Methods (LLM)



Speech input and **noise calibration** is used to capture visitor speech and adjust the recognition threshold based on ambient background noise to improve speech detection.

We used **Speech-to-Text** to convert the recorded speech into text that can be processed by the LLM

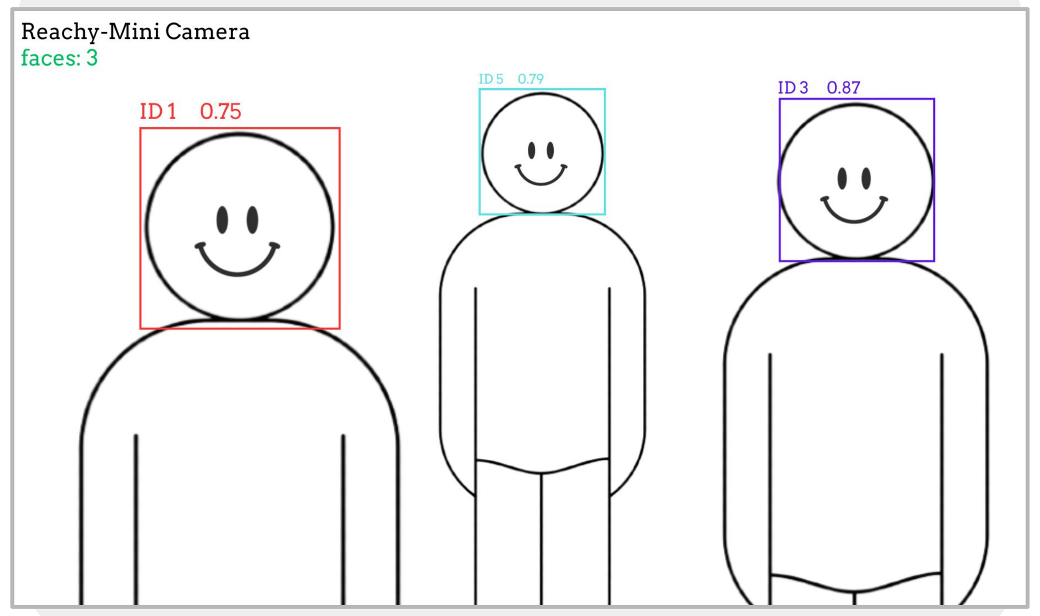
The **LLM** processes the user input to simulate a response that is relevant to the users question

We integrated **prompt engineering** to give Reachy instructions to control its response style and information it should provide.

For the **Text-to-Speech** we used edge_tts to convert the LLM generated response into natural sounding speech

Experimental Methods (Gaze)

Ultralytics YOLO Face Tracker



The **Ultralytics face tracker** uses a YOLOv11n model that assigns unique IDs to track individual faces. Each face is captured with a bounding box and confidence level ranging from 0 to 1.

Person movement translates into robot movement by calculating (x, y) change in position from center of face to change in head angles.

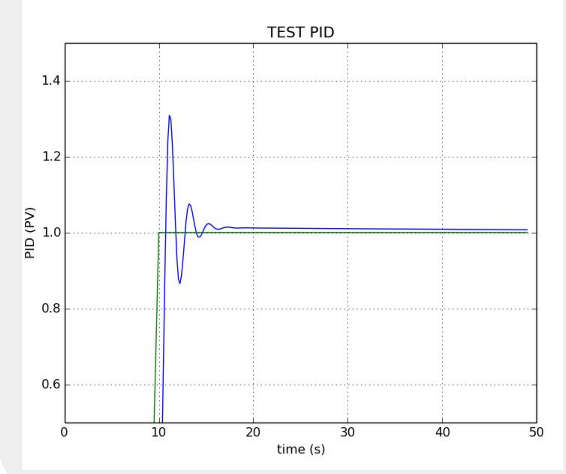
We implemented a **PI (Proportional Integral) Controller** to stabilize and smooth robot movement. The controller combines two methods of minimizing error: **proportional** correction which happens immediately (present) and **integral** error which is accumulated over time (past).

Calculating yaw/pitch angle from change in displacement

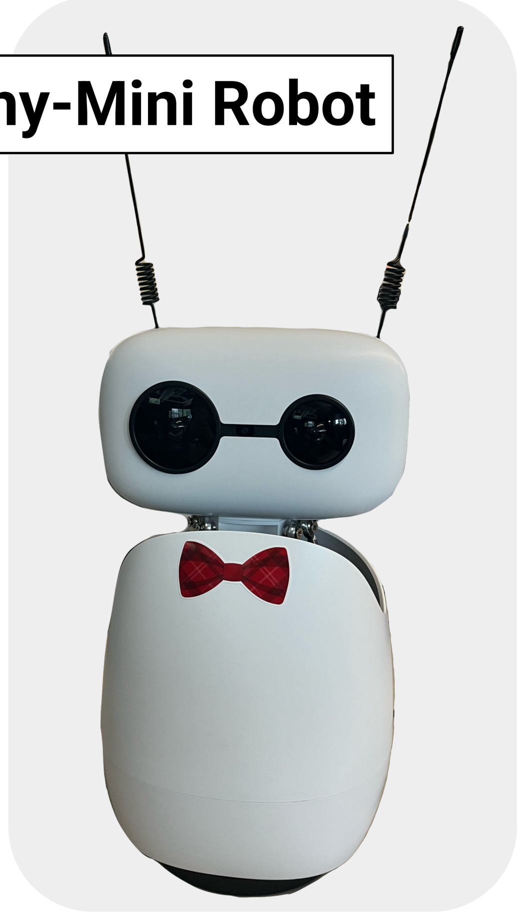
$$\theta = \mathrm{FOV} \times \frac{\Delta x}{\mathrm{Resolution}}$$

change in angle = field of view * (change in position / image resolution)

PI Controller



Reachy-Mini Robot

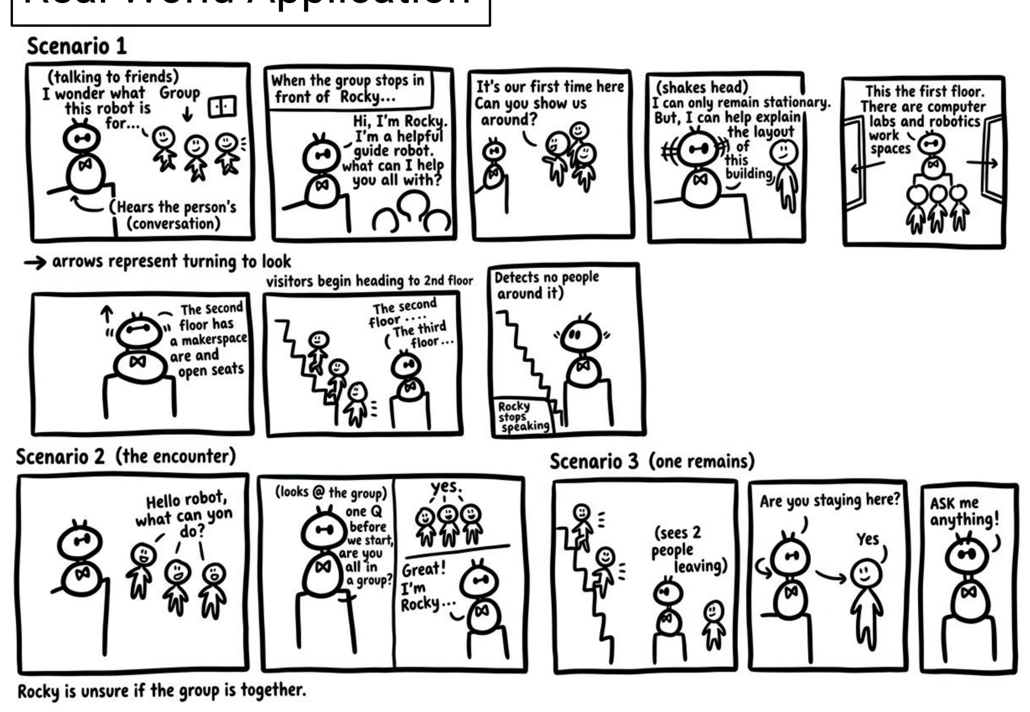


Results

Algorithm 1 Gaze-Mediated Group Membership Signaling

```
Params: min face 80 px; group gap d = 500 px; glance τ_v = 3 s;
    dwell τ ~ U{4, ..., 7} s
while active do                                         ▷ per frame
    F ← tracked faces ≥ 80 px, sorted by ID
    if F = ∅ then continue
    end if
    G ← cluster from lowest ID (grow while gap < d)
    v ← lowest-ID face outside G                        ▷ visitor, if any
    if v unseen this cycle then
        hold gaze on v for τ_v                          ▷ one-shot acknowledge
    else
        gaze at group member g; after τ, advance cyclically
    end if
    PI-track g: head yaw/pitch, body turn if |yaw| > 25°
end while
```

Real World Application



Conclusion and Future Steps

- Create a tabletop robot capable of communicating with dynamic groups of visitors
- Utilize gaze, body movement and dialogue to boost social engagement
- Integrate the speech, LLM, and gaze systems into a unified real time interaction pipeline that allows Reachy-Mini to listen, respond, and direct its attention toward visitors

For the robot to achieve success in the real world as a guide, it is necessary to consider various cases that may happen in real life. We recognize the importance of exposing the robot to diverse social environments in order to appropriately adapt and adjust its behaviors. Following testing, results will be assessed by factors such as duration of interactions and visitor permanence.

Open Questions:
How do different gaze patterns influence visitors' engagement with the conversation?
How can the robot cater to multiple groups simultaneously?

References

Andrist, S., Tan, X. Z., Gleicher, M., & Mutlu, B. (2014). Conversational gaze aversion for humanlike robots. *Proceedings of the 2014 ACM/IEEE International Conference on Human-Robot Interaction*, 25–32. https://doi.org/10.1145/2559636.2559666

Mutlu, B., Shiwa, T., Kanda, T., Ishiguro, H., & Hagita, N. (2009). Footing in human-robot conversations. *Proceedings of the 4th ACM/IEEE International Conference on Human Robot Interaction*, 61–68. https://doi.org/10.1145/1514095.1514109

Tang, F., Hu, M., Fang, Y., Yu, H., Nichols, E., Gomez, R., & Li, G. (2025). Social Robot Haru assisting dynamic group discussion with Autonomous Eye Gaze behavior. *2025 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 6372–6379. https://doi.org/10.1109/iros60139.2025.11245910

Vázquez, M., Carter, E. J., McDorman, B., Forlizzi, J., Steinfeld, A., & Hudson, S. E. (2017). Towards robot autonomy in group conversations. *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*, 42–52. https://doi.org/10.1145/2909824.3020207

Acknowledgements

People And Robot Collaborative Systems (PARCS) Lab
Xiang Zhi Tan, Assistant Professor of Computer Science
Reina Chan, Drake Moore, PhD Students
Center for STEM Education
Claire Duggan, Executive Director
Jennifer Love, Associate Director
Ahmed Othman, Kenneth Santizo & Frances Corcell, YSP Coordinators
Nicolas Fuchs, Program Manager
Mary Howley, Administrative Officer

This work was supported by the National Science Foundation