CPSC 290 Proposal
Advisor: Dr. Marynel Vazquez
Student: Megha Joshi
Fall 2021

Investigation of Robots Influencing Social Context Based on Gaze Behavior

Background:

Perceiving gaze in a social context is one of the most important aspects of interaction. It is a task that seems fairly trivial to humans but can be nuanced in human-robot interaction settings. For example, for screen-based robotic heads, it is known that people might experience a sort of "Mona Lisa effect" when interacting with them. They may perceive gaze from the robot as if it was establishing mutual gaze with them more often than is actually intended or geometrically correct, based on the true gaze direction of the robot. While this effect may seem like a detriment to interactions, perhaps it could be leveraged from an interaction design perspective to balance participation in group conversational settings. This project will build foundational technologies to study this possibility in the future.

To achieve effortless conversation, a robot's screen face should be able to address an individual in a group interaction, prompt conversation, and suggest participation through indirect methods such as gaze. This work is motivated by previous efforts in Human Computer-Interaction, specifically in speech recognition and multi-modal mechanisms for robots to convey gaze direction (figure 1).

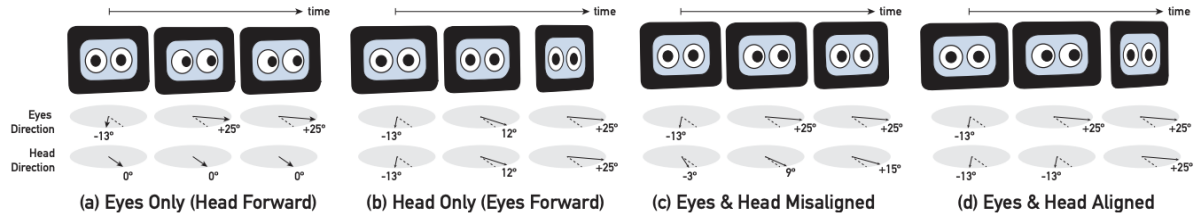

Figure 1: Prior research on development of different gaze behaviors based on eyes direction and head direction. Perceptions of the robot's gaze direction, mutual gaze, width of the robot's cone of direct gaze, and naturalness ratings were studied (1).

There are two primary objectives for this research: the first is to identify who is speaking in a social interaction and for how long they have been speaking, the second is to utilize prior research and code on gaze behaviors and request specific behaviors to be shown to an individual depending upon how long they have been speaking.

Objective One:
To accomplish the first objective, there are two technologies that will be researched. The first is the use of a kinect sensor. This will make it easier to identify when there is someone speaking but not necessarily who is speaking. The second option is to utilize a UMA-8-SP USB mic array (figure 2). This would require configuration with Ubuntu 20.04 to fit Shutter's operating system, as well as creating an association of the audio with the direction of the person.

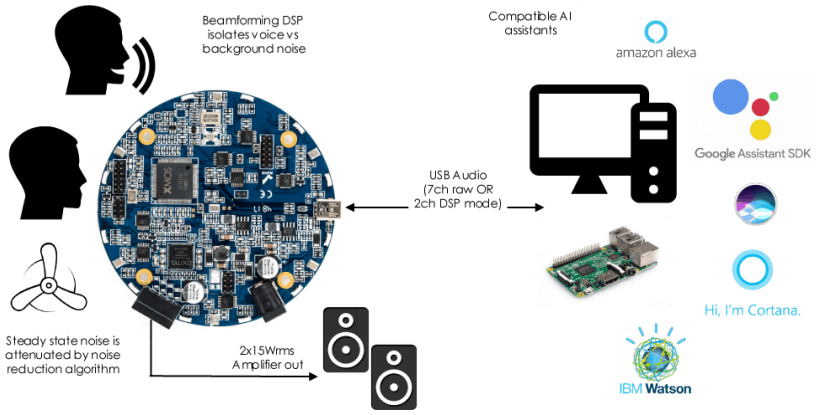

Figure 2: Diagram of the functionality of the UMA-8-SP mic array, explaining how voices are isolated and compatible with common AI assistants (2).

Objective Two:

The second objective consists of developing a program that requests to execute an appropriate gaze behavior depending on the characteristic of that person in a social context. For example, if participant X has had a low speaking ratio compared to the rest of the participants, which would be determined in objective one, then Shutter may direct its gaze to participant X in a way that would urge participation. The programming for this will primarily be done in Python. Worth noting, I will also implement core functionality for this program to customize how the robot coordinates nonverbal communication behaviors (gaze and head motion) as it directs its gaze in a particular direction towards the user(s).

Proposal:

In this project I will work with Dr. Vazquez and another undergrad student to improve Shutter, the Robot Photographer, a custom made social robot developed by the Interactive Machines Group. This research aims to improve the way in which Shutter can perceive social interaction and prompt individuals who have not been speaking enough to do so in an indirect way, specifically through gaze behaviors. I will be the lead of the second objective, and support integration with the first one to create an interactive system that can evaluate speaking turns and adapt in real-time. At the beginning of the semester, I will complete IRB training, and tutorials to understand how the Robot Operating System (ROS) works and to implement my code leveraging its tools and following its conventions. ROS is a middleware used for distributed computation in Shutter. Also, throughout the semester, I will help deploy Shutter to understand how interactions evolve and adapt my code to make it suitable to typical social situations observed with Shutter.

Potential implications:
The addition of this technology would allow a social robot to monitor an observation and take an active role in ensuring participation in a way that seems effortless and indirect to the participants

of that interaction. This work will contribute to the ongoing study of how 2D eye motion can impact a social context by the way in which a user perceives its gaze.

Deliverables:

My deliverables for this project will be focused on three areas:

1. Determining which hardware is most useful for identifying who is speaking and in which direction (research and comparison of Kinect sensor versus UMA-8-SP USB mic array).
2. Developing a program to associate audio direction with a person and keep track of how long they have spoken over time.
3. Implementing different gaze behaviors and developing a program that executes one of these behaviors towards a specific participant in a social context.
4. Preliminary tests on how the gaze behaviors work in public environments, and how well the audio direction detection software can keep track of the state of interactions.

Planning:

I will be meeting on a weekly basis with Dr. Vazquez and the rest of the team working with the Shutter Robot. I will also be communicating via Slack outside of regular meeting times as needed.

Task	Expected Time
IRB and Ros Tutorials	1.5 weeks
Understanding how to relate people's detections with audio directions via geometric reasoning	1.5 weeks
Implementing gaze behaviors	3 weeks
Implementing top level logic to keep track of who's speaking	2 weeks

The remaining time in the semester will be spent doing testing and data collections to check if the implementations were successful.

References:

1) Vázquez, Marynel, et al. "Gaze by Semi-Virtual Robotic Heads: Effects of Eye and Head Motion." *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2020.
2) UMA-8-SP USB mic array. miniDSP. (n.d.). https://www.minidsp.com/products/usb-audio-interface/uma-8-sp-detail?lang=en&gclid=Cj0KCQjwpreJBhDvARIsAF1_BU2ODat9eWmdPd0hm1Volj664nX8WOElZfmgfkOSYWhsyMRjXriiCJcaAgaDEALw_wcB.