Assistive Intent Recognition And Manipulation

I-Chen Jwo, Jiahong Ouyang, Karsh Tharyani, Ting-Che Lin, Yang Yang

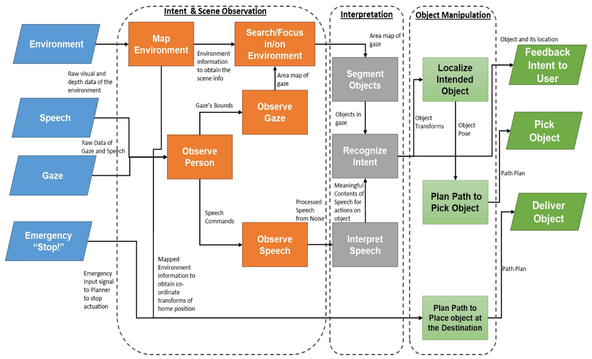
Carnegie Mellon University

THE ROBOTICS INSTITUTE

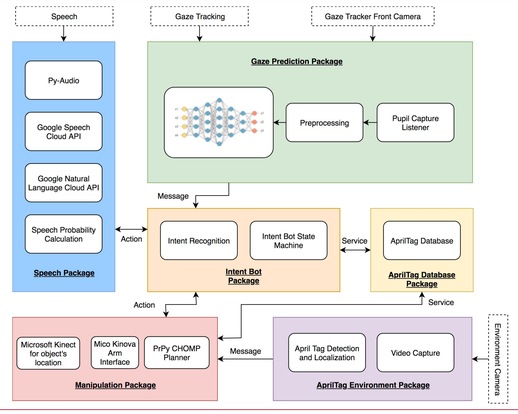
Motivation & Goal

According to statistics, there are more than 19.9 million Americans suffer from upper body limitations. Our goal is to help patients who are struggling with **upper-body dysfunction** to perform everyday tasks like **picking up an object**. We created an assistive robot by constructing **shared autonomy**, which means to predict intention through the **trajectory of the gaze** and **control a manipulator** to complete the task.

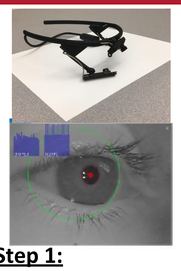
System Architecture



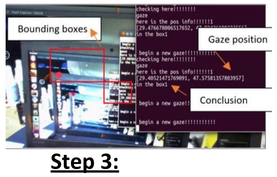
Cyber-physical Architecture



Intent Prediction Subsystem-Gaze Tracker



Hardware: Pupil Labs gaze tracker with two inward cameras to detect pupil and one outward camera to represent the first-ego view.



Step 1:
Adjust the camera to make sure the eye is in focus and all range of the eye movements are visible.

Step 2:
Calibrate the gaze to establish a mapping between pupil and gaze coordinates.

Step 3:
Show the gaze position as the red spot. Publish the position and possibility to analysis.

Use Case Scenario

The Intended Object | User's View and Gaze | Predict the Intention | Planning | *"Yay!"*

Step 1:
User with upper body dysfunction is sitting on a table for lunch and wants the green soup can; she stares at it for 5 - 10 seconds

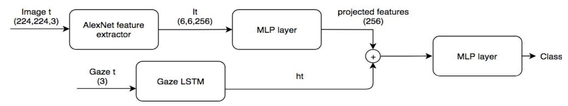
Step 2:
The Intent Bot records and tracks the gaze of the user and sends this intention to the manipulator

Step 3:
The manipulator plans in the cluttered environment to give to the user the desired object

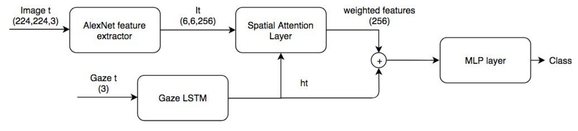
Intent Prediction Subsystem-Network Models

- **Simple Classification Model:** Use the position of gaze to crop the world image. Use AlexNet as a classifier.

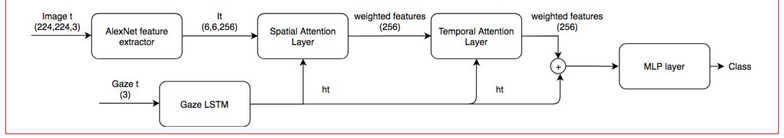
- **LSTM Baseline Model:** Use AlexNet as a feature extractor. Use LSTM to process the gaze sequence. Simply concatenate gaze and image features and feed to the MLP.



- **Spatial Attention Model:** Add spatial attention layer to assign different weights to each region based on the baseline model.



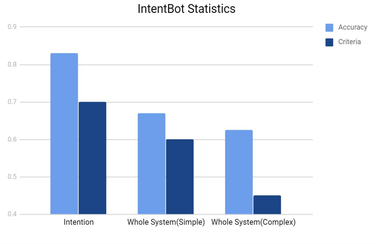
- **Multiple Attention Model:** Add temporal attention layer to assign different weights for each frame based on the spatial attention model.



Testing Procedure & Results

Procedure:

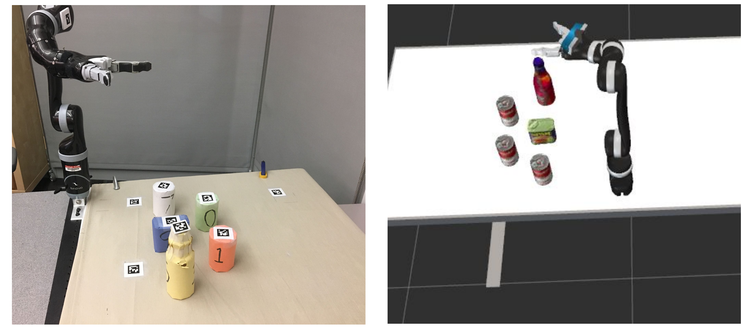
- Re-arrange objects.
- Give the subject the start signal.
- Predict subject's intention by tracking gaze
- Start the manipulator.
- The arm moves the obstacles and reaches for the target.



Manipulation Subsystem

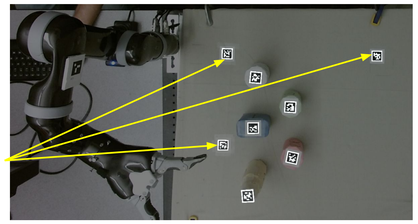
Environment Perception

User's View (Actual World) | Robot's View (Simulation)



Object Localization

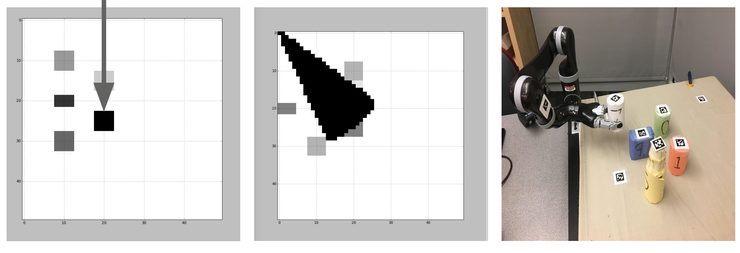
Reference AprilTags



- Fix a Kinect above the table to detect objects and Reference AprilTags.
- Use Reference AprilTags to establish the coordinate and map the position of objects from Kinect coordinate to robot base coordinate.

Planning Pipeline

Intended Target
Object's tag ID



Step 1:
Make a Footprint map of the objects

Step 2:
Find obstacle-objects in the Target's way; their new locations

Step 3:
Action! Plan to intended object using TSR and CBiRRT

Future work

- Deep learning based signature detection approaches
- Onboard data processing and adaptive planning
- Please visit our website for more information