Language Guided Manipulation

Group 5

Sashank Visweshwaran
Hanah Zachariah
Shreya Patki

Preliminary Work

Goal: To be able to manipulate the environment using natural language commands.

- Habitat AI home robot: Open Vocabulary Mobile Manipulation

  Explores environment, finds an object, finds a receptacle, places the object

- RT-X Models: RT-2

  Large vision language model that outputs robot actions based on natural language tokens

Creating a natural language model and integrating it with an existing system, to be able to pick up the correct object based on text input provided by the user.

Experiments

- Habitat AI home robot implementation for language based manipulation of the robot

- Embodied QA implementation through navigation using House3D dataset

  Agent is asked a question, navigates its environment to gather information for answering it.

- Inverse kinematic implementation from scratch using inverse kinematic implementation for grasping, along with a basic language model for determining the object based on text input.

- Dope: Pretrained on household objects dataset. Used for getting object pose based on the camera input.

Implementation Details

- Utilized the Scene Replica implementation for getting the pose of objects and grasping them.
- Created our own dataset for different types of instructions for every object. (direct and indirect instructions)
- Trained a SVC model using this dataset, with an accuracy of 86%
- Converted the learning model and the vectorizer to .pkl files to include it in the python script.
- Designed a script to get the text command from the user and pass it to the language model to predict the correct object.
- Performed grasping using Scene Replica's model grasping logic.

Demo

Future Work

- Dropping the objects in a specific location
- Using a different algorithm/methodology for getting the pose of an object
- Using different techniques for grasping
- Expanding dataset and make it applicable to unseen objects (using unsupervised learning techniques)

References

Scene Replica: https://github.com/IRVLUTD/SceneReplica

Habitat AI Home Robot: https://aihabitat.org/

RT-X: https://robotics-transformer-x.github.io/

Embodied QA: https://github.com/facebookresearch/EmbodiedQA