P U M A S

National Autonomous University of Mexico
Tokyo Information Design Professional University

Daniel Vanegas[1], Itzel González[1], Miguel García[1], Ruth Moreno[1], Héctor Robles[1], Marco Negrete[1], Jesús Savage[1], Jorge Solano[1], Luis Contreras[2] and Hiroyuki Okada[2]

[1]Faculty of Engineering, National Autonomous University of Mexico, México

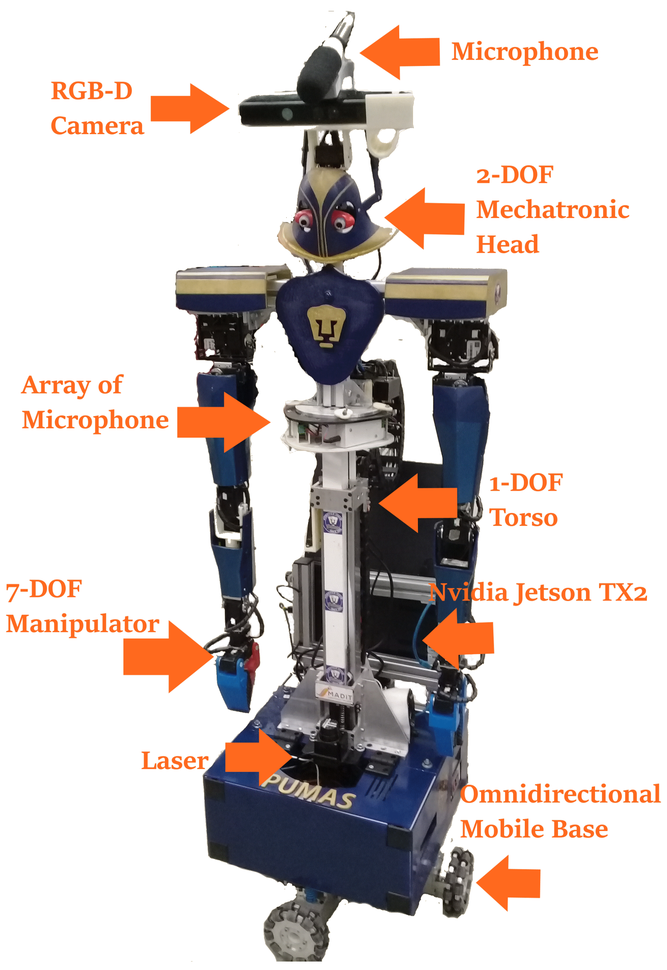
[2]Tokyo Information Design Professional University, Japan

Introduction

Justina is a Service Robot designed for operation in domestic environments. It's friendly and smart design eases the interaction with humans and their environment, while taking advantage of lowcost components available in the market. Justina is part of the initiative of getting robots closer to people to enhance quality of life. However, the primary goal of the project is to encourage and promote science and engineering among students, giving them the opportunity of apply their knowledge in realworld applications using high-end technology

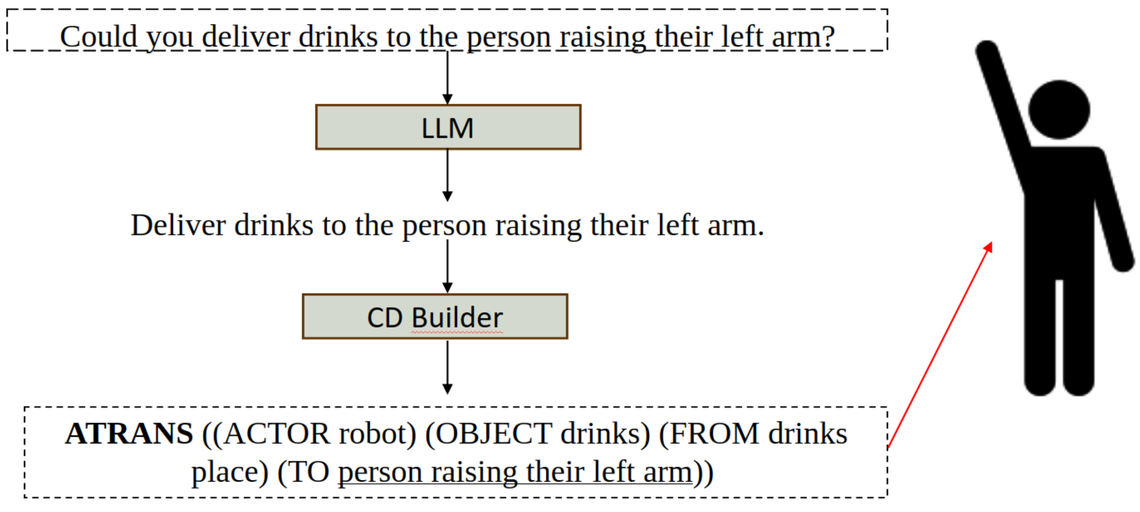
Hardware

- **Mobile base:** Omnidirectional.
- **Manipulators:** 2 x 7-DOF anthropomorphic arms with 10 Dynamixel servomotors each.
- **Head:** 2-DOF (Pan and tilt) built with Dynamixel servomotors.
- **Torso:** 1-DOF through a worm screw and a configuration of gears.
- **RGB-D Camera:** Microsoft's Kinect sensor.
- **Microphone:** Rode NTG2 directional microphone.
- **Laser:** Hokuyo rangefinder URG-04LX-UG0.



Towards General Purpose Service Robots

We propose a framework to process natural language commands for general-purpose service robots where the robot should perform an arbitrary spoken command requested by a non-expert operator. Our system uses a Natural Language Processing (NLP) parser and a Conceptual Dependency (CD) builder to create CD structures that an expert system can employ to generate a global plan that the robot will execute. An extra simplification step using Large Language Models (LLM).



This project was supported by UNAM-DGAPA under grant IT102424

Software

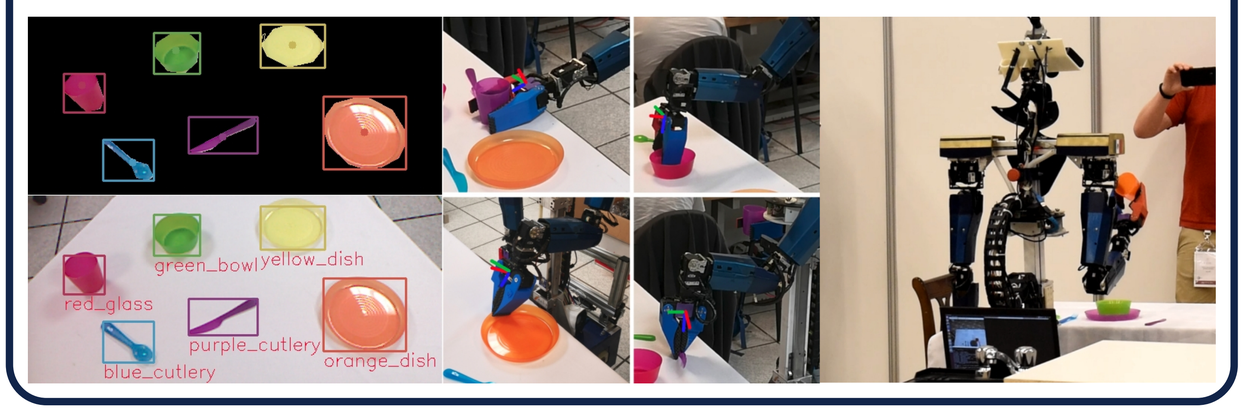
This robot is based on the ViRbot architecture for autonomous mobilerobot operation, which provides a platform for the design and developmentof software for general purpose service robots.The ViRbot architecture defines a human-robot interaction interface madeof three main layers

- Actions planning
  - Space-state search and hierarchical task networks.
  - Using clips how inference engine
  - Hardwired solutions.
  - Task accomplishment Supervisor unit.
- Motions Planning
  - Navigation using Kinect
  - Free/occupied space clusterization.
  - Ocupancy grid + A*
  - Potential field for close-range navigation.
- Vision
  - Histogram Disparity, YOLO and MaskRCNN for Object Recognition
  - Haar Cascades and Facenet for face recognition and detection
  - Openpose for gesture recognition
- Middleware: ROS Kinetic; Blackboard
- Speech
  - Speech Synthesis Loquendo
  - Microsoft Speech Recognition SAPI 5.3
  - Sphinx
  - Natural Language Understanding: Conceptual Dependency and NLP parser.

Intelligent flat-and-textureless object manipulation

- Raw color segmentation, refined by morphological and hull operations
- Principal component analysis on segmented masks
- Eigenvectors and eigenvalues determine each category dish, cutlery, bowl and glass

For manipulation, the robot end-effector roll, pitch and yaw angles are determined according to the object class to be manipulated.



Gesture recognition

We use the openpose running on the Nvidia Jetson TX2 for human's key points detection, our algorithm make a cluster using the closest points in the point cloud corresponding for each key point in order to get a skeleton 3D representation. This representation is used as the input of a Hidden Markov Model (HMM) to determinate the most probably gesture.