Development of intelligent systems (RInS)

Task 3: FooDeRo

Danijel Skočaj
University of Ljubljana
Faculty of Computer and Information Science

Academic year: 2021/22

2022 Final task

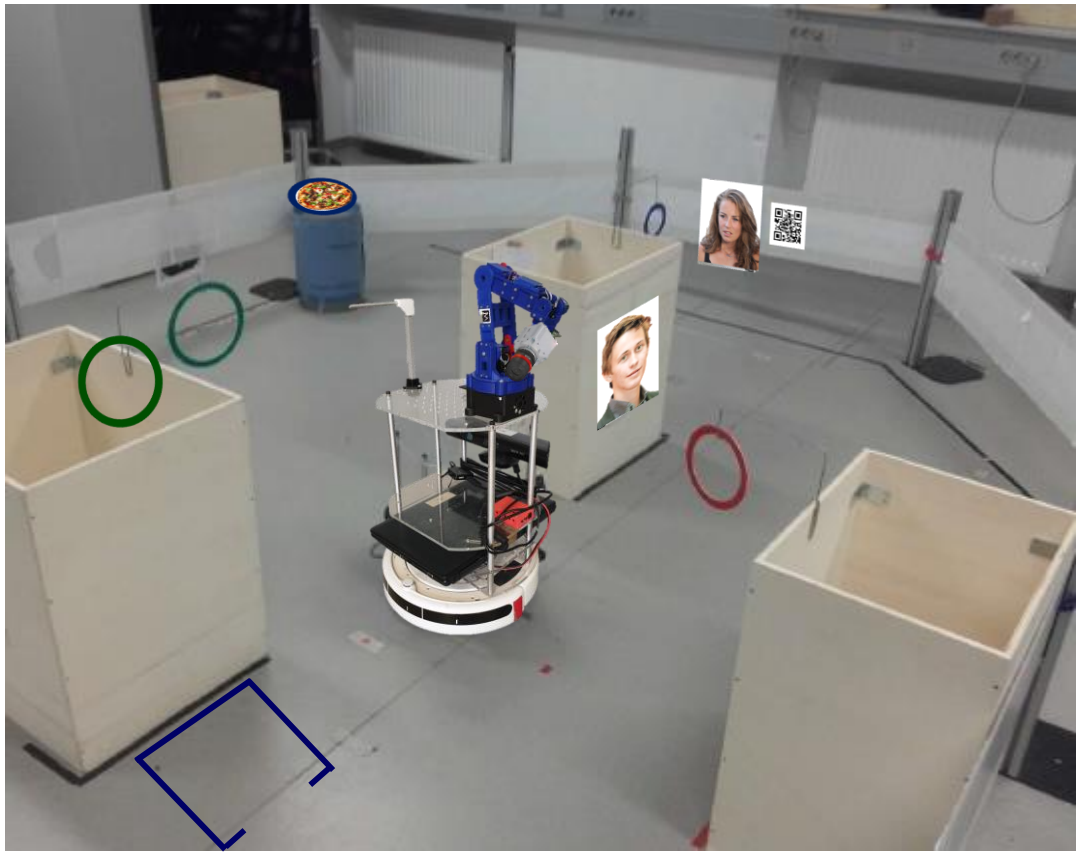
FooDeRo

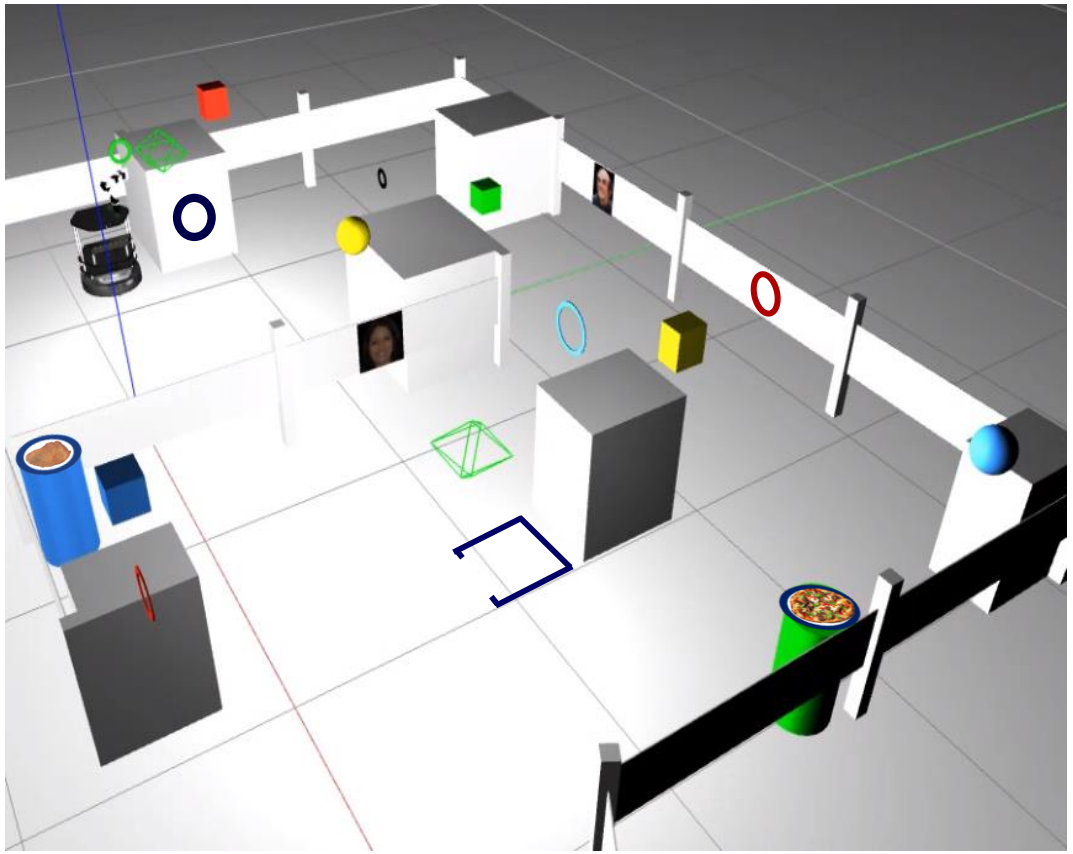
Food Delivery Robot

2022 Final task

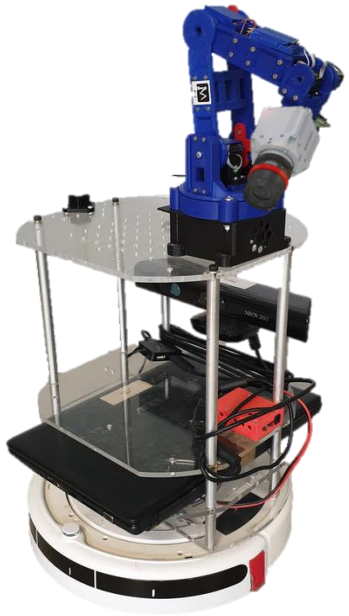
- Setup:
  - „Small city“ scene (fenced area).
  - Several persons (faces) in the scene.
  - Three „restaurants“ (cylinders) of different colours.
  - Four parking slots marked with rings of different sizes and colours.
- Goal:
  - Deliver the (virtual) food to the persons.
- Task:
  - Find all persons in the city.
  - Find all restaurants and look on top of them which food do they serve.
  - Park in the starting parking slot (marked with the green ring).
  - Accept orders from the link given in the QR code.
  - Collect the food and (virtually) bring it to the corresponding persons.
  - Talk to them briefly.

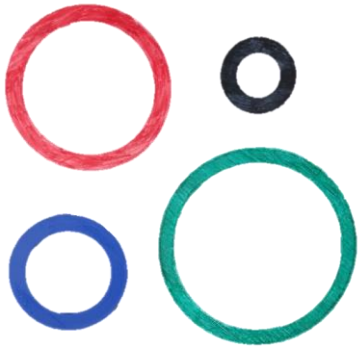
2022 Final task



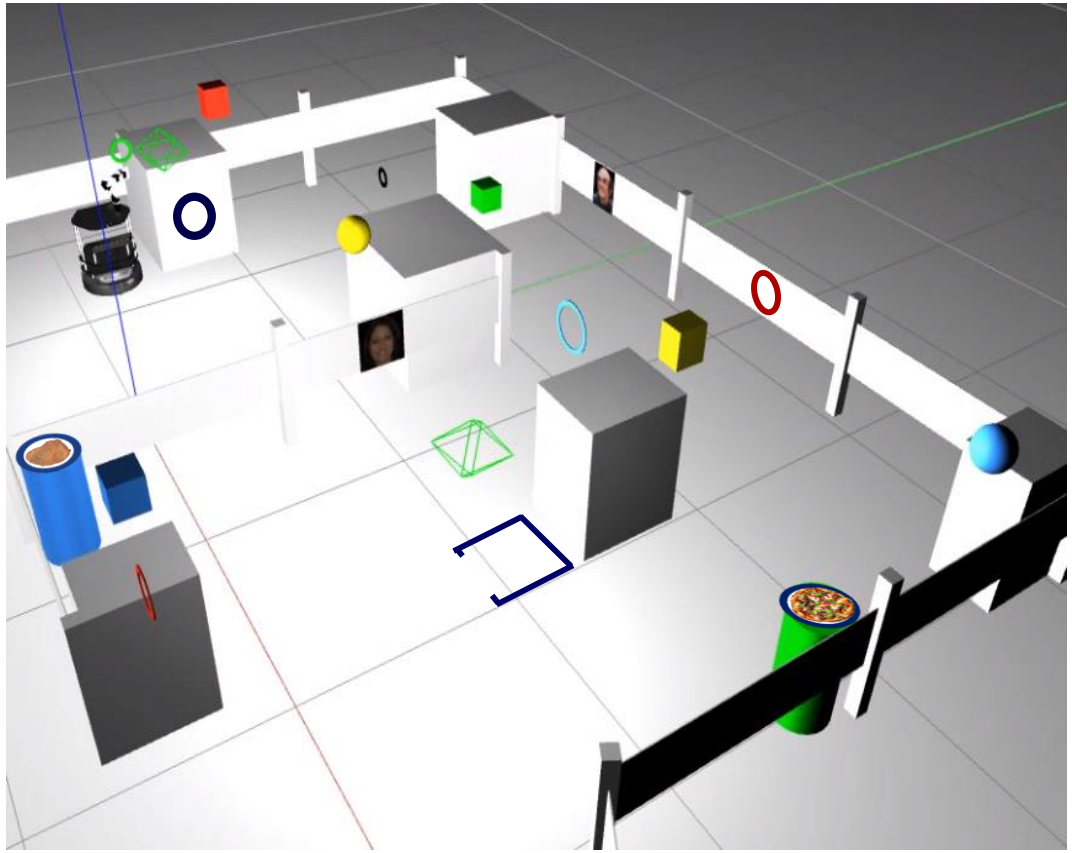


2022 Final task





Simulated scene



- There will be faces and other objects on the walls, including printed rings
- There might be different objects in the scene at the same height as 3D rings
- Rings will not be positioned exactly above the parking place

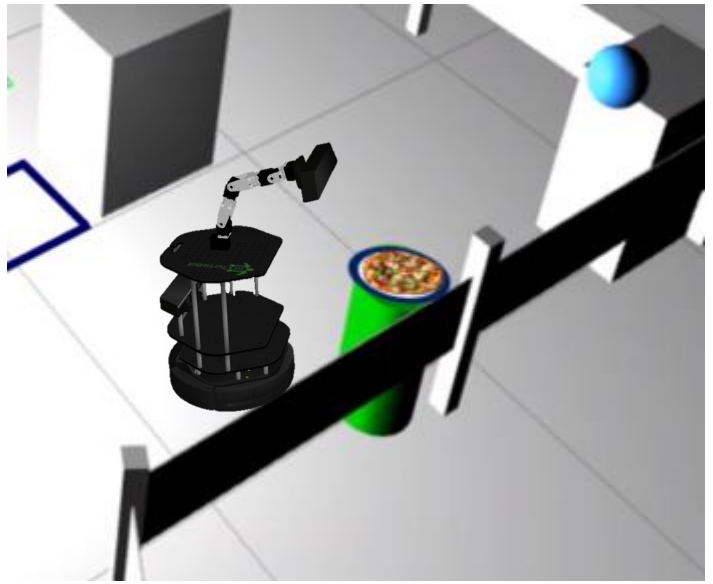
More precisely

- Find and recognise all persons
  - Detect a person, recognise his identity and remember his position
  - The number of persons present is not known in advance
  - (Preferably) implement the autonomous exploration of space
- Find all restaurants (cylinders)
  - Approach them and look on top of them
- Recognise food
  - Detect the plate (a circle) and recognise the food using image classification
  - Or use object detection directly
- Go to the starting parking slot
  - Detect the green ring
  - Park in the corresponding parking slot
- Accept orders
  - Read the QR code
  - Read the information from the link encoded in the QR code
- Pick up and deliver the food
  - Approach the restaurants (cylinders) serving the food specified at the link
  - Approach the persons specified at the link
  - Minimize the distance travelled needed to fulfil this task
- Have a dialogue with the person
  - Virtually „deliver the food“ and have a simple dialogue with the person

Even more precisely

- Find all persons and remember their positions
  - Done for Task 1
  - The number of persons present is not known in advance
- Recognise persons
  - Face recognition
  - A limited number of faces and their identity given in advance
  - Learning classifier or KNN on embeddings
  - Utilise/integrate/develop/train any face/object recognition algorithm

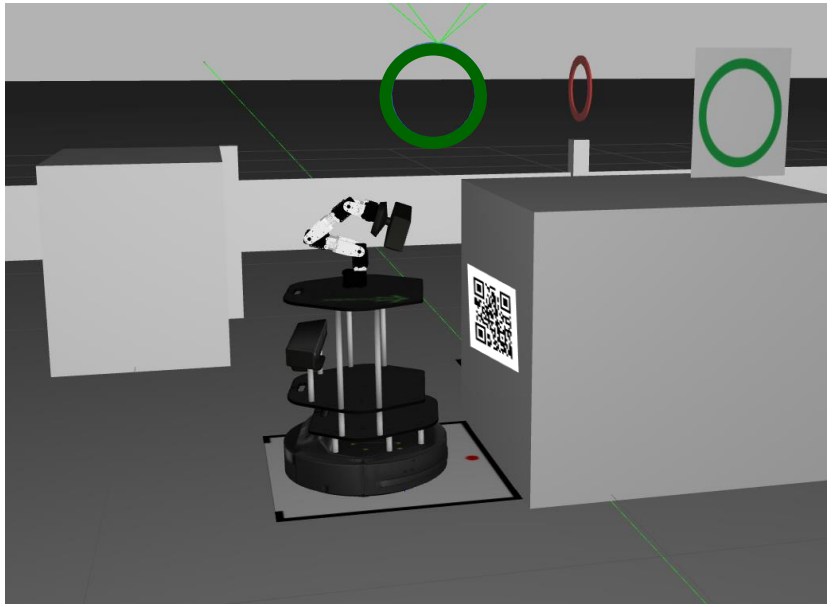
- Find all restaurants (cylinders)
  - Done for Task 2
  - There will be three cylinders in the scene
- Approach them and look on top of them
  - Approach close enough
  - Use the second camera to look on top of the cylinders



Even more precisely

- Recognise food
  - Detect the plate (a circle) and recognise the food using image classification
  - Or use object detection directly
  - Utilise/integrate/develop/train any object detection or recognition algorithm
  - The food will always be positioned on the circle plate
  - Several types of the food will be available
  - Examples of the food images will be given in advance

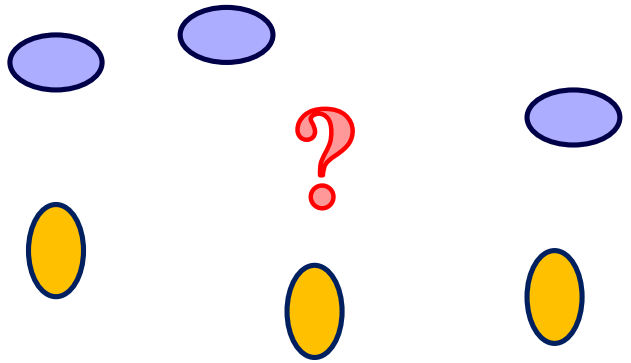
- Go to the starting parking slot
  - Detect the green ring
  - Park in the corresponding parking slot
  - Done for Task 2
  - The parking slot will not be positioned exactly under the ring



Even more precisely

- Accept orders
  - Detect the QR code on the wall in the front of the parking slot
  - Read the information from the link encoded in the QR code
  - Simple format: `name food`

- Pick up and deliver the food
  - Approach the restaurants (cylinders) serving the food specified at the link
    - Indicate that you are picking up the food („I will take one pizza!“)
  - Approach the person specified at the link
    - Approach close enough and look into him
  - Minimize the distance travelled needed to fulfil this task
  - Plan the path accordingly
  - Every person will ask for one type of food
  - Maximum of three persons will place an order



Even more precisely

- Have a simple dialogue with the person
  - Virtually „deliver the food“
  - R: „Here is your pizza.“
  - H: „Thank you.“
  - Receive the payment from the person
  - R: „Will you pay by cash or credit card?“
  - Wave with the robot arm accordingly:
    - H: „With cash.“ – wave to the left
    - H: „With credit card.“ – wave to the right
  - R: „How satisfied were you with the service on the scale from 1 to 5?“
  - H: „5“.
  - R: „Thank you and good by.“
- Simple speech synthesis
- Simple speech recognition, limited vocabulary
  1. Automatic speech recognition
  2. Text input (also as a backup)

Shortcuts

- You may not implement all the functionalities (for a lower grade)

- You can find out some information using the input to a text dialogue box:
  - Which person ordered which food
  - What is the colour of the cylinder that serves a specific food
  - Which payment method will a specific person use
  - How satisfied with the service a person was

- You can therefore skip the following tasks
  - Parking
  - QR code reading
  - Food detection and recognition
  - Speech recognition

- You can also use predefined goals instead of autonomous space exploration

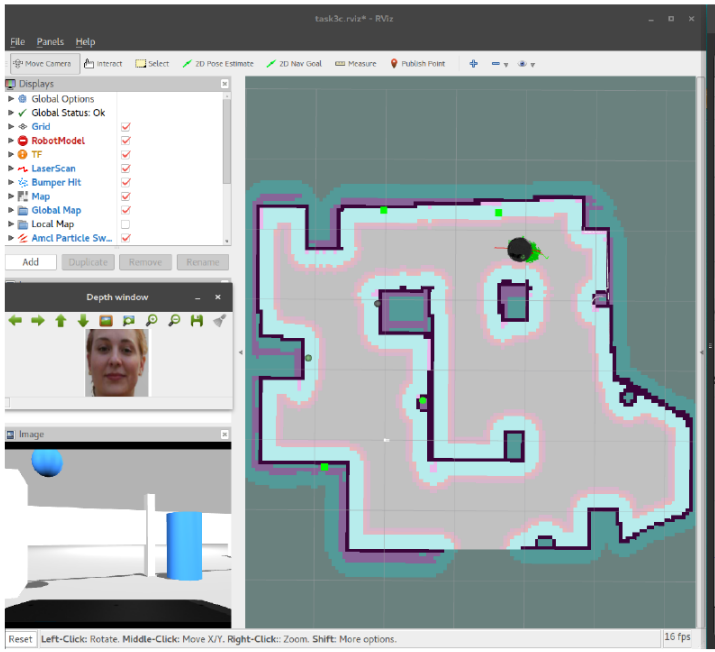
Demonstration

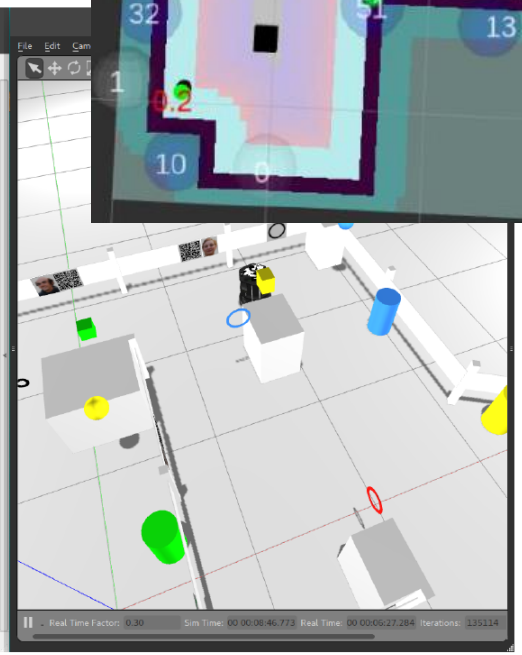
- Demonstrate what is going on in the robot
- Visualize in RViz:
  - Locations of detected cylinders, rings, faces
  - Recognised colours, food
  - Navigation goals, path plans
  - Current sensor readings (images, Lidar)
- Show dialogue in a separate window
- Show the reasoning process
- Show also the current environment in Gazebo

R: Here is your pizza.
P: Thank you.

R: Will you pay with cash or credit card?
H: With cash.

R: How would you grade the service?
H: With 5!
R: 5! Thank you!





Tasks

- System setup
  - Running ROS
  - Tele-operating TurtleBot
- Autonomous navigation
  - Autonomous control of the mobile platform
  - Acquiring images and 3D information
  - Simultaneous mapping and localization (SLAM)
  - Path planning, obstacle avoidance, approaching
  - Advanced fine manoeuvring and parking
  - Intelligent navigation and exploration of space
- Advanced perception and cognitive capabilities
  - Detection of faces, circles, 3D rings, 3D cylinders
  - Recognition of faces, food, digits, colour
  - Basic manipulation and visual servoing
  - Speech synthesis, speech recognition, dialogue processing, reading QR codes
  - Belief maintenance, reasoning, planning

Task 1

Task 2

Task 3

Integrate everything into a robust coherent system

Evaluation protocol

- The evaluation course will be set up in advance
  - The main setup will not change
- The teams will be allowed to build the map in advance
- The faces, cylinders, food, parking places and the rings will be positioned on the day of the evaluation
  - Faces, and the size and colours of the cylinders and rings, as well as types of food are known in advance
- The robot has to operate completely autonomously
  - only the initial positioning is allowed
  - (and the optional answering by typing the text)
- The robot can start at any position

- Every team will have allocated 15-20 minutes to show the performance of the robot

Requirements

- System setup
  - Running ROS
  - Tele-operating TurtleBot
- Autonomous navigation
  - Autonomous control of the mobile platform
  - Acquiring images and 3D information
  - Simultaneous mapping and localization (SLAM)
  - Path planning, obstacle avoidance, approaching
  - Advanced fine manoeuvring and parking
  - Intelligent navigation and exploration of space
- Advanced perception and cognitive capabilities
  - Detection of faces, circles, 3D rings, 3D cylinders
  - Recognition of faces, food, digits, colour
  - Basic manipulation and visual servoing
  - Speech synthesis, speech recognition, dialogue processing, reading QR codes
  - Belief maintenance, reasoning, planning

For 6
For + max. 2
For + max. 2

Grading

- Must do:
  - Face detection (2 pts)
  - Ring detection (2 pts)
  - Cylinder detection (2 pts)
  - Colour recognition (1 pt)
  - Face recognition (1pt)
  - Approaching cylinders (1 pt)
  - Approaching faces (1 pt)
- Should do:
  - Taking images of food (1 pt)
  - Food recognition (2 pt)
  - Parking (2 pts)
  - QR-code detection (1 pt)
  - Taking orders from the link (1 pt)
  - Optimal path planning (1 pt)
  - Auton. space exploration (2 pts)
  - Dialogue with ASR (2 pts)
  - Weaving with manipulator (1pt)

- Performance evaluation
  - Navigation (1 pt)
  - Reasoning (1 pt)
  - Visualisation (1 pt)
  - Robustness (1 pt)
  - Relative speed (1 pt)
  - Overall impression (2 pts)

- Points:
  - Must do: 10
  - Should do: 13
  - Performance: 7
  - Total: 30

Task 3 goals

- The main goals of the third task and evaluation are:
  - to navigate the robot around
  - to detect faces in 2D
  - to detect objects (rings and cylinders) in 3D
  - to learn and recognize colours
  - to learn and recognize faces
  - to learn and recognize food images
  - to do simple reasoning
  - to do simple dialogue processing
  - to plan adequate actions
  - to fine manoeuvre the robot
  - to do simple mobile manipulation
  - **to integrate all functionalities into a coherent system**