Development of intelligent systems (RInS)

Task 3: Robo da Vinci

Danijel Skočaj
University of Ljubljana
Faculty of Computer and Information Science

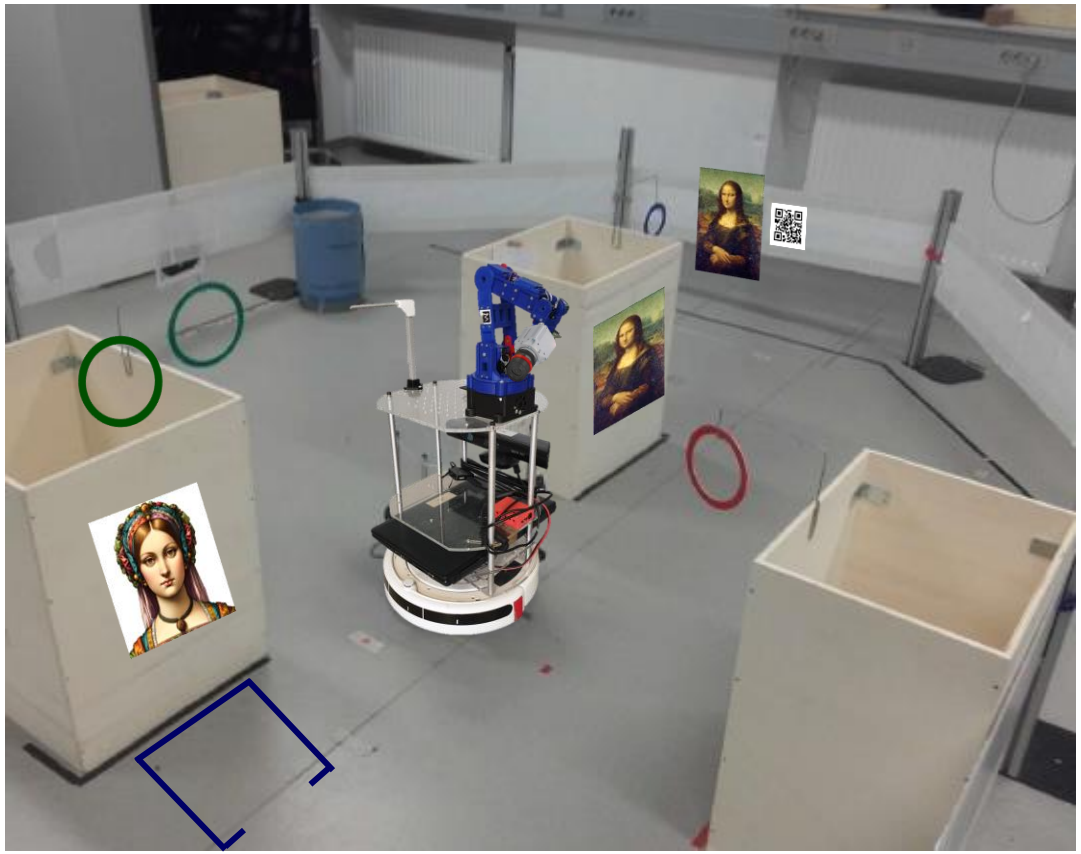
Academic year: 2023/24

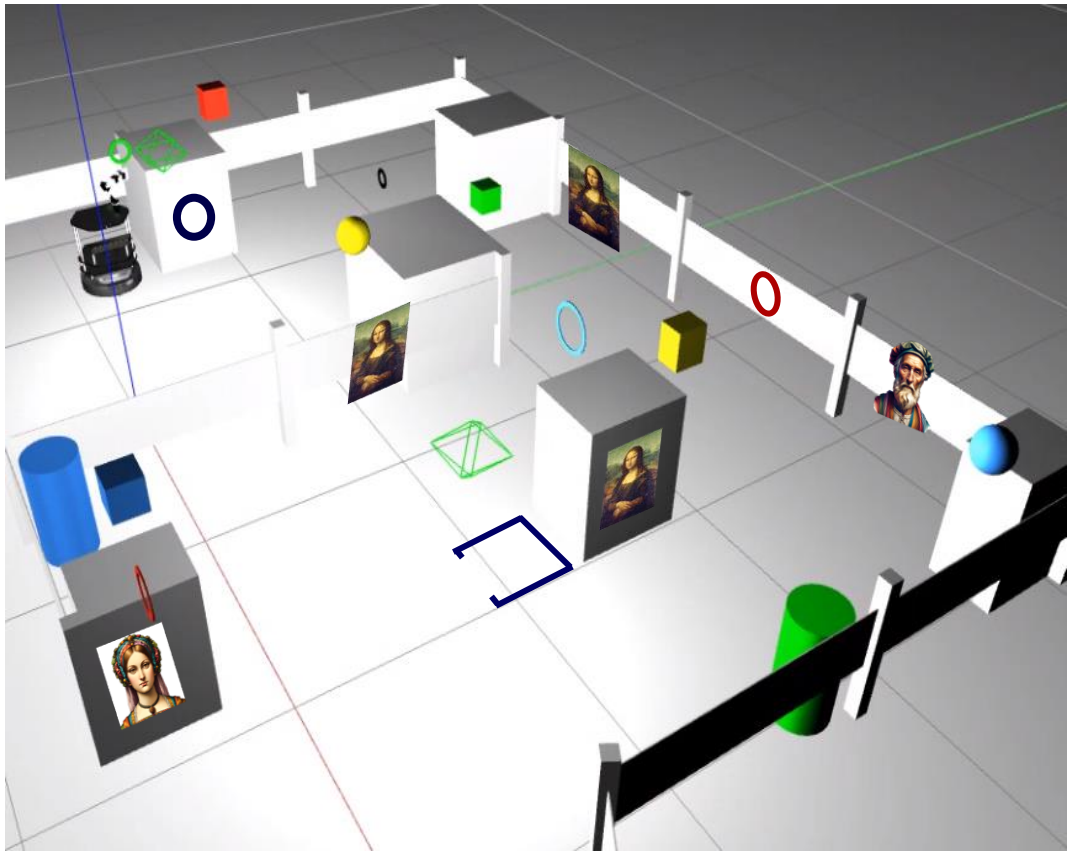

Robo da Vinci

2023 Final task

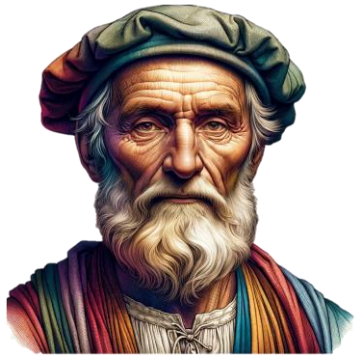
- Setup:
  - „A small Renaissance town “ scene (fenced area).
  - Several people (faces) in the scene.
  - Several paintings of Mona Lisa in the scene.
  - Four „art galeries“ (cylinders) of different colours.
  - Four parking slots marked with rings of different sizes and colours.
- Goal:
  - Find the genuine Mona Lisa painting.
- Task:
  - Find all the people in the city.
  - Talk to people to find out where a photo of the genuine Mona Lisa painting can be found.
  - Park in the corresponding parking space and obtain the photo of the genuine Mona Lisa (from the link provided in the QR code).
  - Find all the paintings of Mona Lisa.
  - Analyse the paintings and determine which one is genuine.
  - Go to all the forged paintings and display evidence on an image as to why each is forged.
  - Go to the genuine Mona Lisa and point at it.

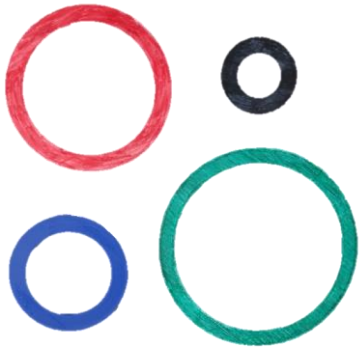
2023 Final task



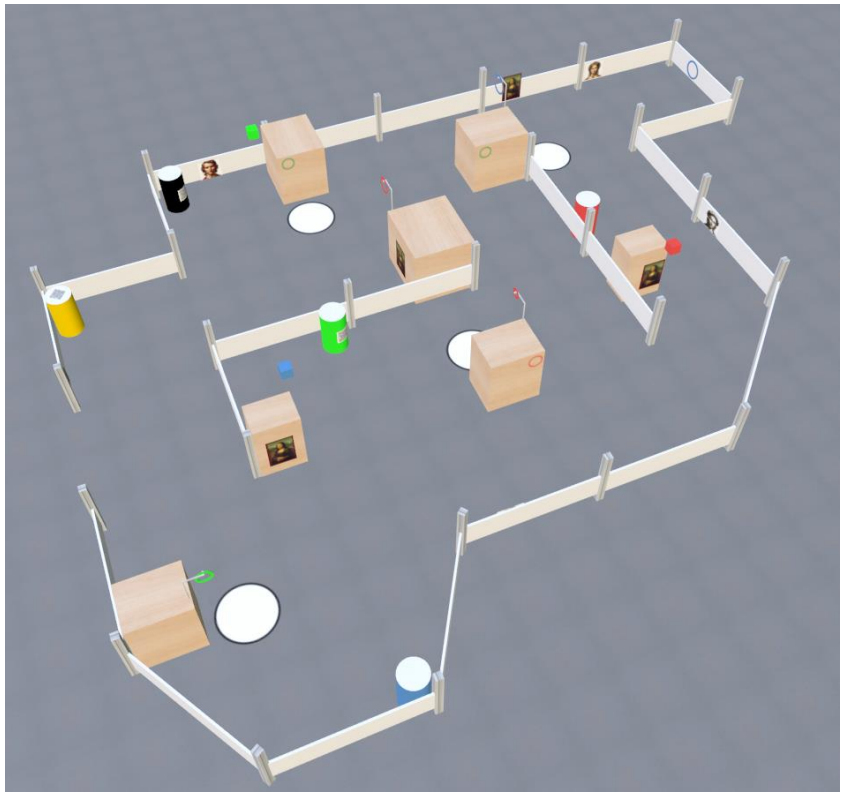


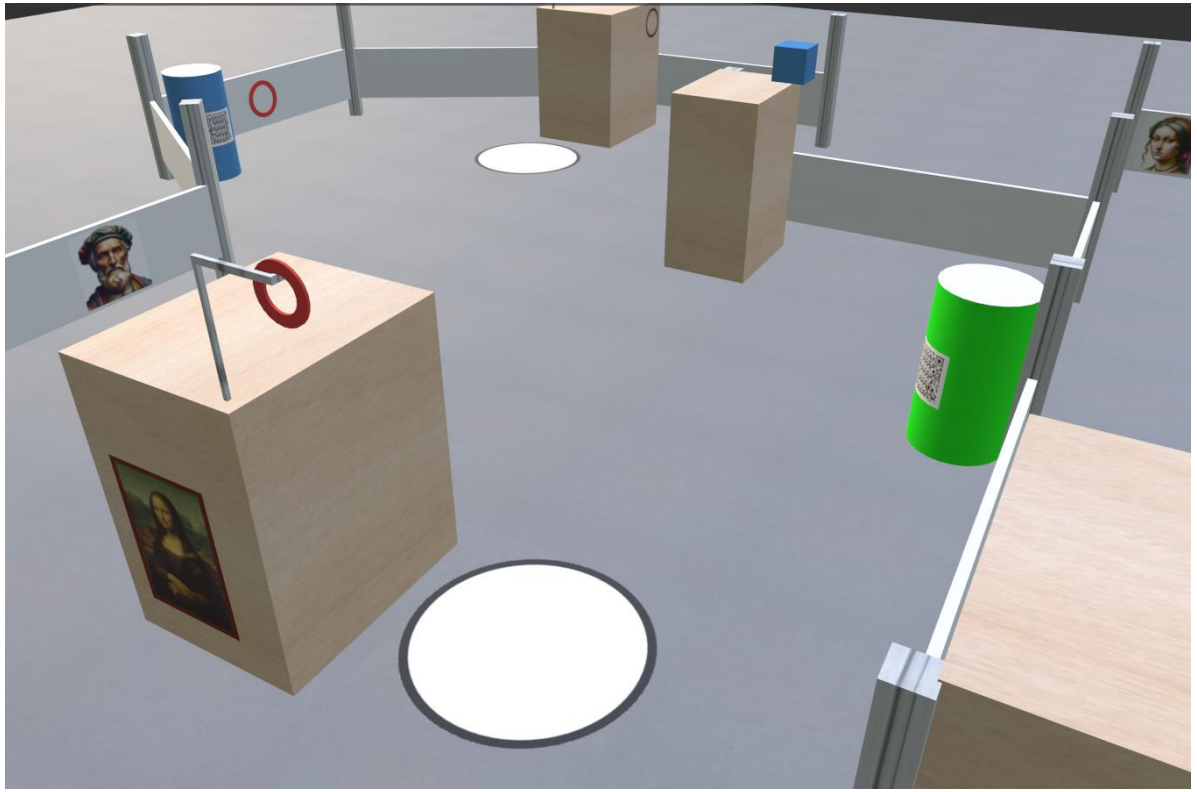
2023 Final task





Simulated scene

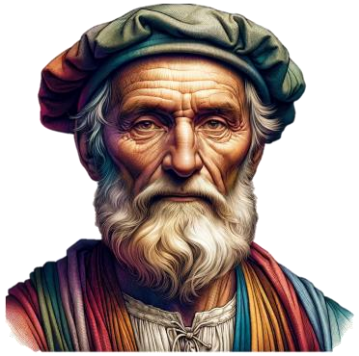




- There will be faces, paintings, and other objects on the walls, including printed rings
- There might be different objects in the scene at the same height as 3D rings
- Rings will not be positioned exactly above the parking place

Even more precisely

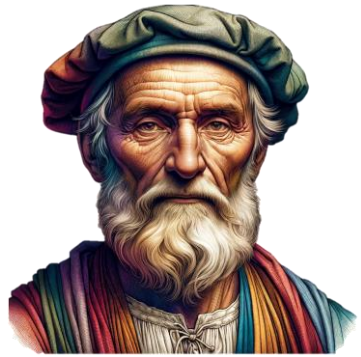
- Find all persons
  - Done for Task 1
  - The number of persons present is not known in advance
  - Differentiate between persons and Mona Lisa paintings
  - (Preferably) implement the autonomous exploration of space
- Find all Mona Lisa paintings
  - By detecting faces, and recognising Mona Lisa
  - By detecting the paintings directly
  - By detecting the frames of paintings





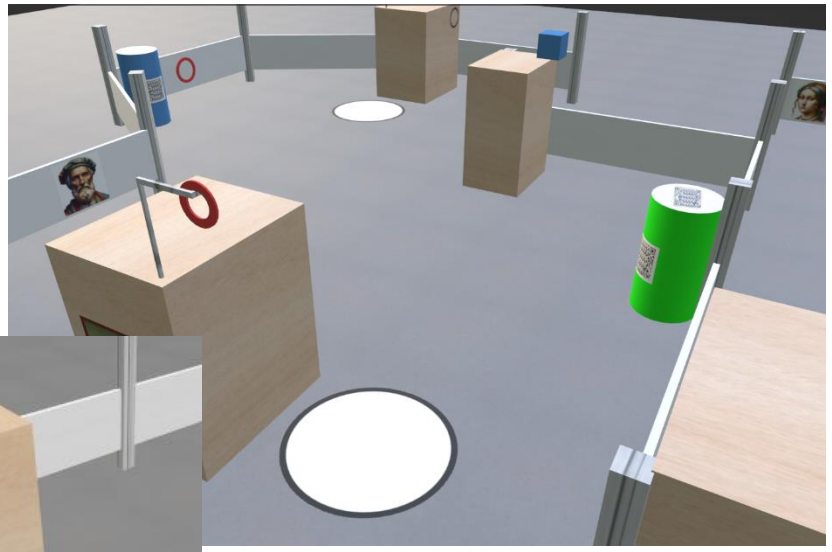
Even more precisely

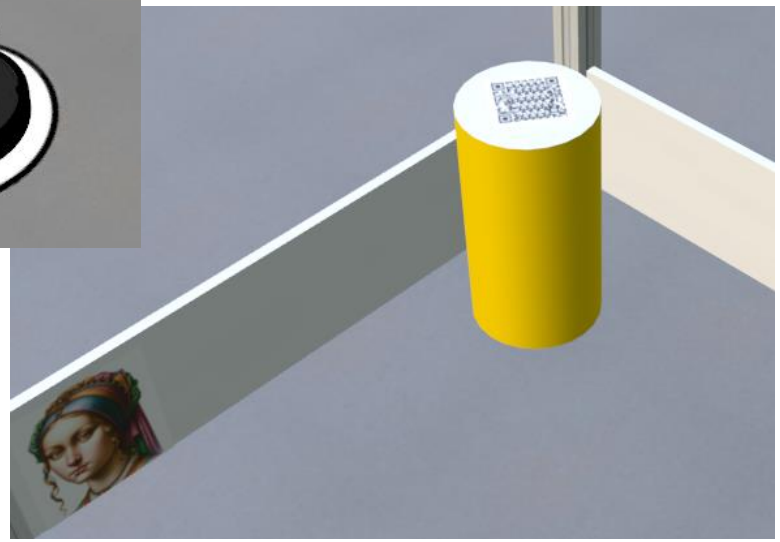
- Engage in a dialogue with people to determine where a photo of the genuine Mona Lisa painting can be located.
  - Approach a person
  - Ask them about the color of the cylinders that mark the correct parking position.
  - Implement a simple dialogue:
    - R: „Do you know where the info about the MonaLisa photo can be found?“
    - H1: „No, I don‘t.“
    - R: „Do you know where I should look for the Mona Lisa photo?“
    - H2: „Park under the green or blue ring and look around.“
    - R: „OK. Thank you.“
  - Two of the people will give a useful hint; the others know nothing.
  - Each hint will always include two parking places (different pairs ;-)
  - The robot should not speak to paintings!
  - Simple speech synthesis
  - Simple speech recognition, limited vocabulary
    - Automatic speech recognition
    - (Or text input as a backup)



Even more precisely

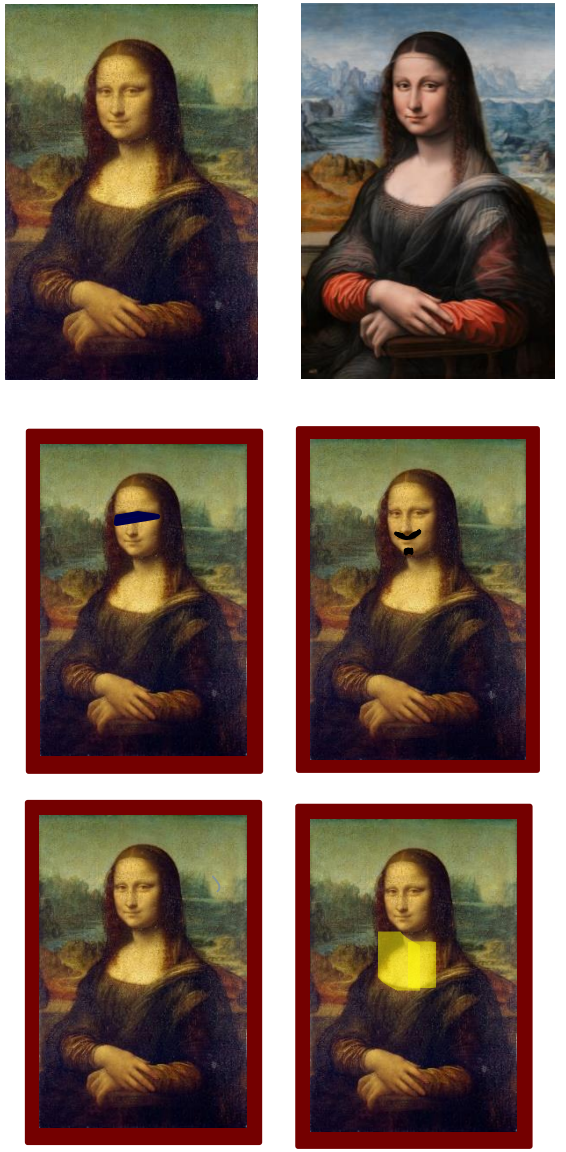
- Find the parking place and park
  - Detect the rings
  - Recognise the colour of the rings to determine the correct ones
  - Park in the parking slot situated beneath the corresponding ring
  - (Done for Task 2)
  - The parking slot will not be located exactly under the ring
- Obtain a photo of the Mona Lisa painting
  - Look around and detect the cylinder (done for Task 2)
  - "If there is a cylinder within a radius of 2 meters, detect and read the QR code on top of it.
  - Retrieve the image from the link encoded in the QR code — it is a photo of the genuine Mona Lisa painting.





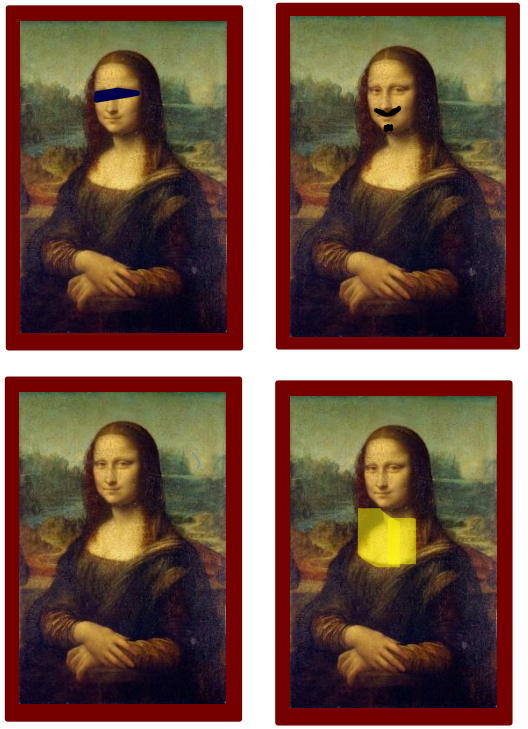
Even more precisely

- Create/train a surface anomaly detector
  - Based on PCA, AE, VAE, normalising flow,difusion models, discriminative, reconstructive methods, etc.
  - Two images, of a genuine and another Mona Lisa will be given in advance
  - Train an anomaly detection model for each of it
  - Two levels of difficulty:
    - Severe degradations of the painting (glasses, etc.)
    - Mild degradations (more difficult to detect)
  - Take a number of training images and augment them
    - In Gazebo, in simulation, as you want
    - Taken from the frontal view or from side views as well
    - Vary the illumination conditions etc.
  - Use or design a method for anomaly detection
  - Train the model
- Evaluate the anomaly detection model
  - Create also test images
  - Evaluate the model under different environments
- Train the classifier to select the correct model
  - Could be very simple choice between two images
  - Image recognition, retrieval, similarity check
- Can be done (almost) completely outside ROS and Gazebo



Even more precisely

- Find the genuine Mona Lisa painting
  - Compare the detected paintings with the photo of the genuine one
    - i.e., use the model that corrsponds to the photo
  - Utilise the surface anomaly detector
    - Approach to every detected painting
    - Analyse it using the trained anomaly detector
    - Determine wheter the painting is genuine or not
  - If the painting is a forgery:
    - highlight the anomalous regions on the image
    - have the robot verbally express a negative observation
  - If the painting is the genuine Mona Lisa:
    - direct the robot to point at it with its manipulator and wave
    - have the robot making a positive verbal affirmation
  - When the robot finds the genuine Mona Lisa painting, it should then stop.



Shortcuts

- You may not implement all the functionalities (for a lower grade)
- You may not implement the autonomous exploration of space
  - and can use fixed goals instead.
- There will be a QR code containing all necessary information
  - You can position the robot in front of it at the beginning if you decide to use this option
- The robot can skip the dialogue with a person
  - and read the colours of the rings written in the QR code
- The robot may skip parking
  - and use the colour of the cylinder written in the QR code and search for the cylinder directly
- The robot may skip detecting the QR code on the cylinder
  - and read the image of the genuine Mona Lisa directly
- The QR code will therefore include all necessary information in a simple format:

```
ring green
ring blue
cylinder green
image http://box.vicos.si/…
```

Demonstration

- Demonstrate what is going on in the robot
- Visualize in RViz:
  - Locations of detected faces, painting, rings, cylinders, rings
  - Recognised colours
  - Navigation goals, path plans
  - Current sensor readings (images, Lidar)
  - Show live strem from both cameras
- Show the dialogue in a separate window
- Show the reasoning process
- Show also the current environment in Gazebo
- Information in the console does not suffice!
- Use also sound to demonstrate what is going on

R: Do you know where the info about the Mona Lisa photo can be found?
H: No, I don‘t.
R: Do you know where I should look for Mona Lisa photo?
H: Park under the green or blue ring and look around.
R: OK. 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, surface defects
  - Recognition of colour, faces
  - Basic manipuation 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, paintnings, parking places and the rings will be positioned on the day of the evaluation
  - The size and colours of the cylinders and rings 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

- The evaluation will take place in the last week of the semester

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, surface defects
  - Recognition of colour, faces
  - Basic manipuation 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)
  - Painting detection (2 pts)
  - Ring detection (2 pts)
  - Colour recognition (1 pt)
  - Approaching faces (1 pt)
  - Speech synthesis (1 pt)
  - QR code reading (1 pt)
- Should do:
  - Parking (2 pts)
  - Cylinder detection (2 pts)
  - Severe anomaly detection (2 pts)
  - Also mild anomaly detect. (2 pts)
  - 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 pts)
  - Relative speed (1 pts)
  - 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 detect surface anomalies
  - 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**