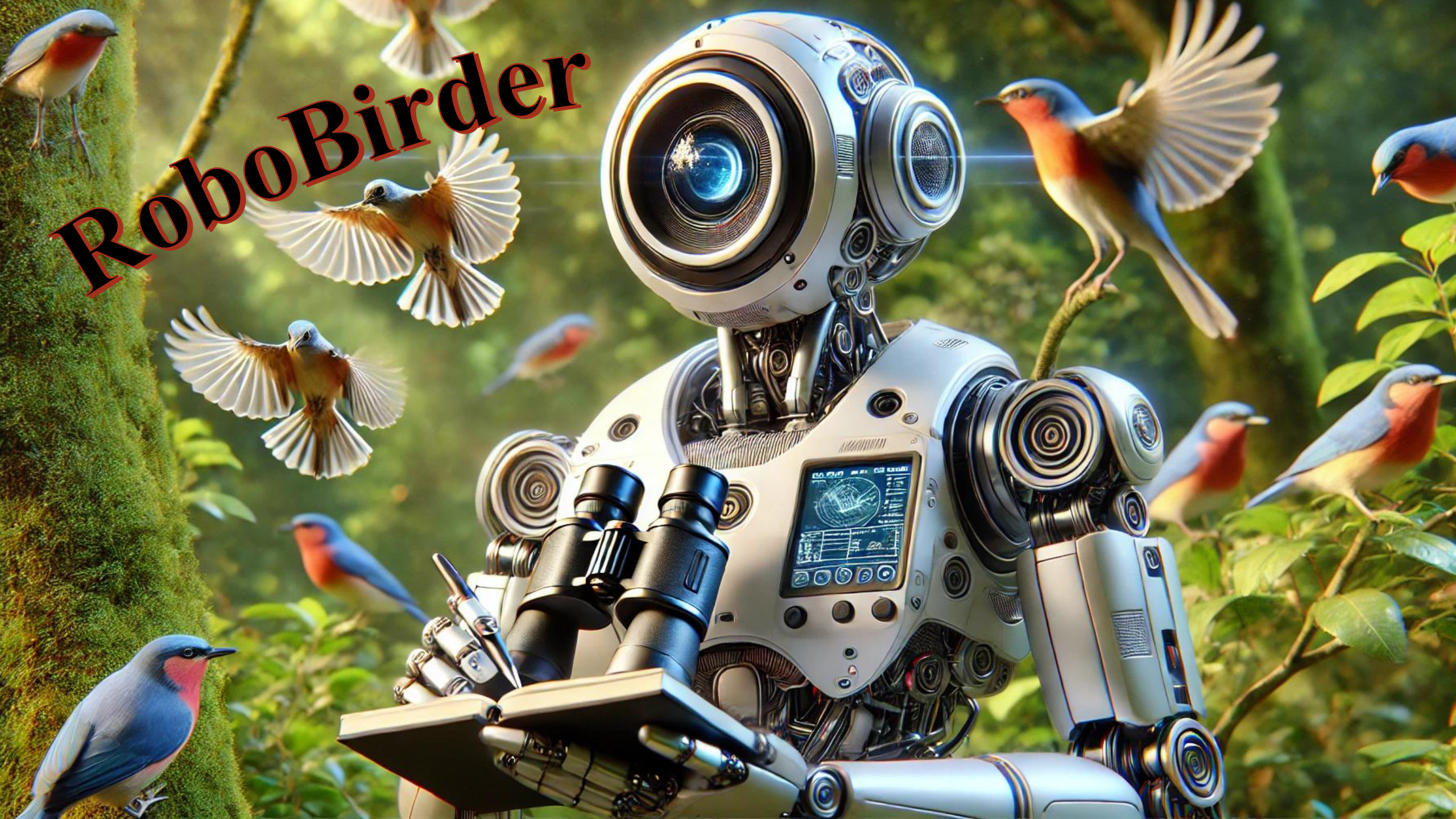
Development of intelligent systems (RInS)

Task 2: RoboBirder

Danijel Skočaj
University of Ljubljana
Faculty of Computer and Information Science

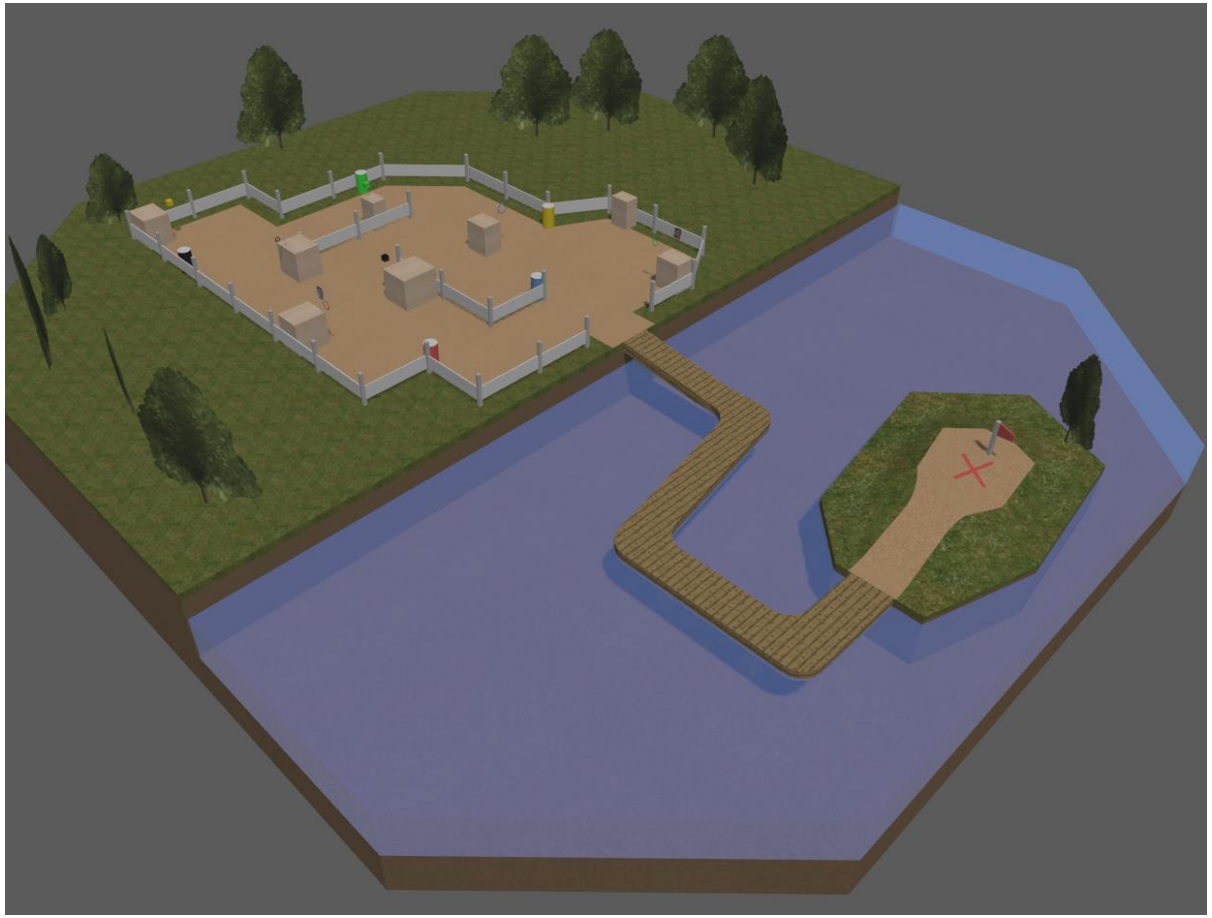
Academic year: 2024/25


RoboBirder

2025 Final task - RoboBirder

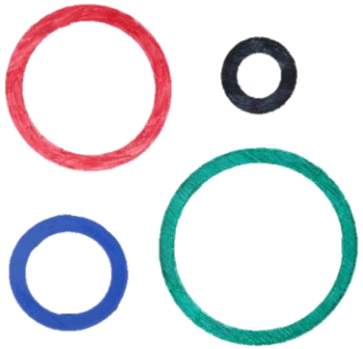
- Setup:
  - „A park“ scene (fenced area).
  - Pathways, bridges, grass area, river.
  - Several persons (faces) in the scene.
  - Several birds in the scene standing on the rings of different colours.
- Goal:
  - Find and recognise all the birds.
- Task:
  - Navigate safely around the park and look around.
  - Find all birds in the park.
  - Approach every bird and take a photo with hi-res camera.
  - Recognise the bird and say its name.
  - Find all persons in the park.
  - Talk to persons to find out which are their preferred birds.
  - Tell the persons where their preferred bird is and what is the colour of the ring it is standing on.
  - Create a catalogue of the detected birds.
  - Pay attention to stay on a pathway and not to drive in the grass or water, and deliver the catalogue to the end point.

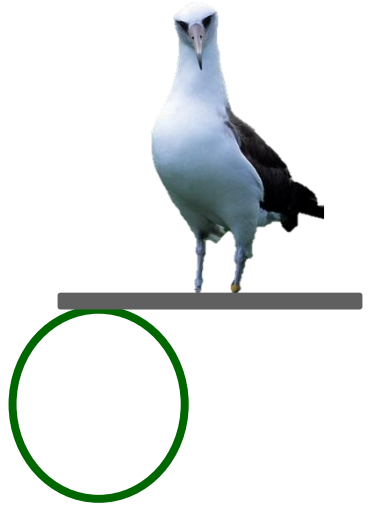
Simulated scene



- There will be faces, birds, and other objects on the walls, including printed circles
- There might be different objects in the scene at the same height as 3D rings

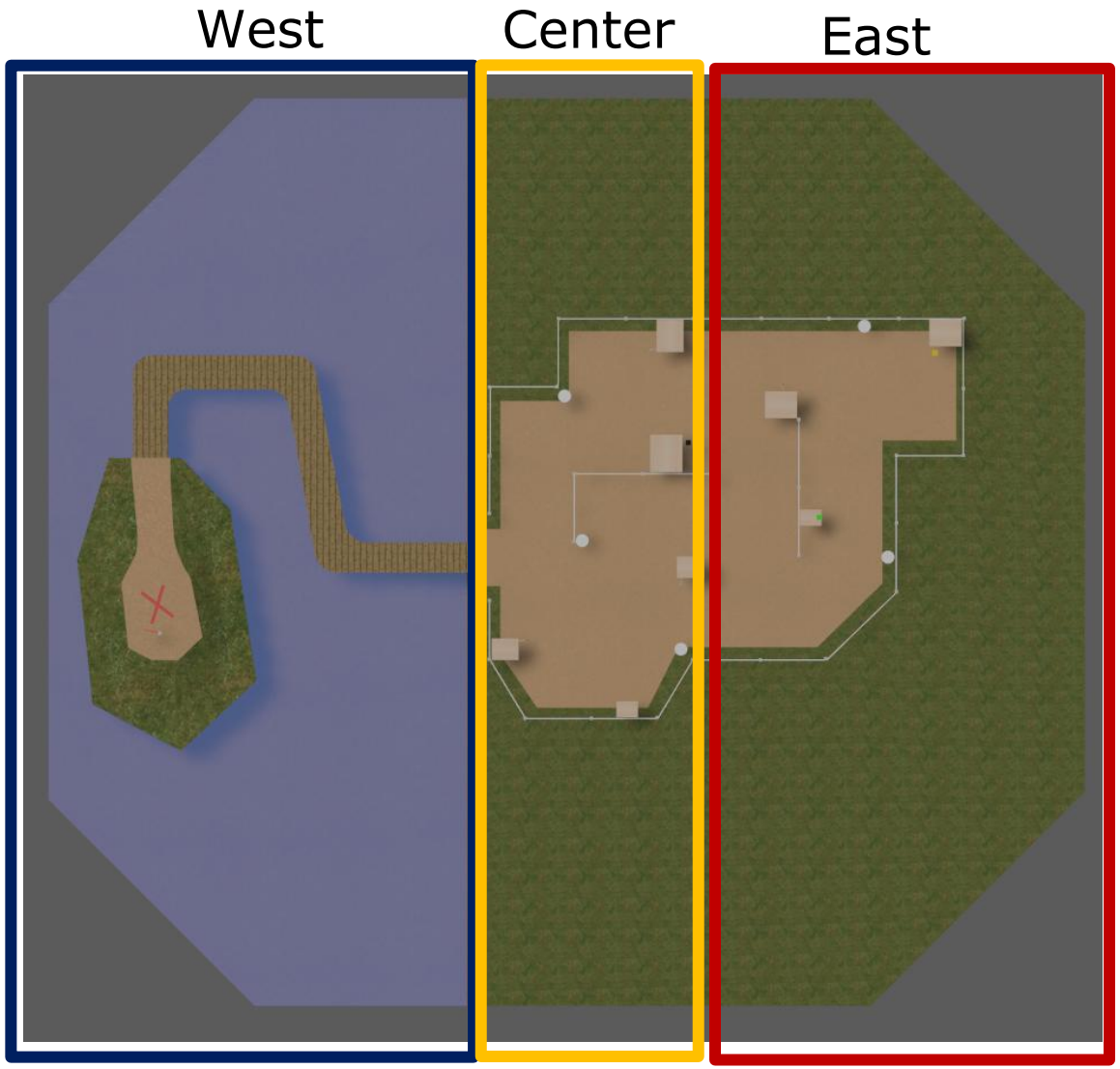
2025 Final task





Navigation

- Explore the park
  - By using a predetermined path or
  - Preferably implement the autonomous exploration of space
- Avoid traveling on the water and on the grass
  - Use the second camera pointed downwards to see what is in front of the robot
  - Do not enter green and blue regions
  - Or add the prohibited regions in the map
  - Image-based navigation only at the end of the task
  - Finish at the end point
- Divide the park in three regions:
  - East, Center, West



Perception

- Ground segmentation for image based navigation
  - At the end of the task
- Find all birds
  - All the birds will be sitting close to the rings
    - A flat bird cut out from an image with a bird
  - Detect the birds directly or
  - Detect the rings and look for birds above them
  - Detect the rings and recognise the colours of the rings
  - Recognise the bird species
  - Approach the bird and say its species name
  - Point the camera toward the bird, take the photo of the bird and store it, along its species name
- Find all persons
  - The number of persons present is not known in advance
  - Recognise the gender of the person
  - Approach the person and engage in a dialogue with him or her

Perception and dialogue

- Engage in a dialogue with people to determine which birds do they like most.
  - Approach a person
  - Recognise the person‘s gender
  - Ask the person about the favourite bird
  - Women always tell their favourite bird directly
  - Men might reconsider their choice
    - The robot has to ask them twice or more
    - When the same bird is named twice, the bird is considered to be the favourite one
    - The person can also confirm the favourite bird with „yes.“

Sample dialogue

- R: „Hi woman, which is your favourite bird?“
- W1: „The robin.“
- R: „Well there is one robin in the center of the park sitting o a green ring.“

- R: „Hi woman, which is your favourite bird?“
- W2: „The hawk.“
- R: „There is one hawk in the central part of the park sitting on a yellow ring.“

- R: „Hi man, which is your favourite bird?“
- M1: „My favorite bird is the swallow.“
- R: „Are you sure?“
- M1: „Well, maybe I like eagles more?“
- R: „OK, the eagle then. Are you sure?“
- M1: „Ah, not really. The swallow is still the best“.
- R: „OK. The swallow then. There is one in the west part of the park sitting on a red ring.“

- R: „Hi man, which is your favourite bird?“
- M2: „The robin.“
- R: „Are you sure?“
- M2: „Yes, I‘m sure.“
- R: „There is one robin in the center of the park sitting on a green ring.“

- R: „Hi man, which is your favourite bird?“
- M2: „My favorite bird is the sparrow.“
- R: „Are you sure?“
- M2: „Well, maybe I like storks more.“
- R: „OK, the stork then. Are you sure?“
- M2: „Ah, not really. I prefer cuckoos more.“.
- R: „Are you sure now?“
- M2: „Yes.“
- R: „OK. The cuckoo then. There is one in the east part of the park sitting on a blue ring.“

Bird classification

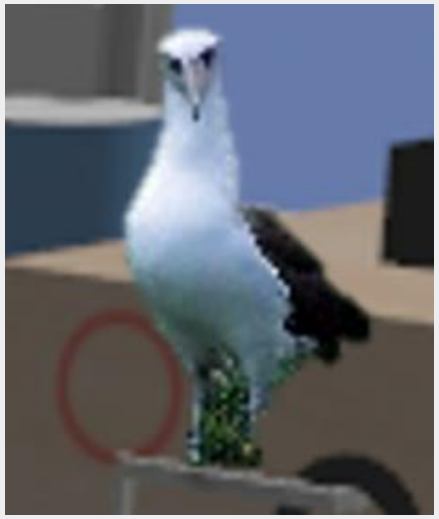
- The Caltech-UCSD Birds-200-2011 Dataset

Bird classification

Bird catalogue

- A PDF file with all detected birds

American goldfinch



Laysan albatross

Fish crow

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
  - the first and the exit point in the park can always be handcoded
- The robot can skip the dialogue with a person
  - and read the favourite bird in the QR code next to the person
- The robot does not need to classify the faces according to gender
  - you assume that all the persons answer the right answer right away
- The robot does not need to point at the bird with the camera
  - it takes a photo of a bird from the distance
- The robot does not need to create a bird catalogue
  - it recognizes the birds, but it does not create a pdf file
- The robot could not avoid walking on the water and grass
  - i.e., it can finish prematurely

Demonstration

- Demonstrate what is going on in the robot
- Visualize in RViz:
  - Locations of detected faces, rings, birds
  - Recognised colours, birds
  - 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: „Hi man, which is your favourite bird?“
M: „The robin.“
R: „Are you sure?“
M: „Yes, I‘m sure.“
R: „There is one robin in the center of the park sitting on a green ring.“

Tasks

- System setup
  - Running ROS — Task 1r
  - Tele-operating TurtleBot — Task 1s
- Autonomous navigation — Task 2
  - Autonomous control of the mobile platform
  - Acquiring images and 3D information
  - Simultaneous mapping and localization (SLAM)
  - Path planning, obstacle avoidance
  - Advanced fine manoeuvring
  - Basic mobile manipulation
  - Intelligent navigation and exploration of space
- Advanced perception and cognitive capabilities
  - Detection of faces, 3D rings, birds
  - Recognition of birds, ring colours, and gender
  - Segmentation of the ground
  - Speech synthesis, speech recognition, dialogue processing (reading QR codes)
  - Belief maintenance, reasoning, planning

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, rings, and birds will be positioned on the day of the evaluation
  - The size and colours of the 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 starts at the starting 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

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
  - Advanced fine manoeuvring
  - Basic mobile manipulation
  - Intelligent navigation and exploration of space
- Advanced perception and cognitive capabilities
  - Detection of faces, 3D rings, birds
  - Recognition of birds, ring colours, and gender
  - Segmentation of the ground
  - 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)
  - Colour recognition (1 pt)
  - Approaching faces (1 pt)
  - Speech synthesis (1 pt)
  - Bird detection (2 pts)
  - Bird recognition (3 pts)
  - Approaching birds (1 pt)
- Should do:
  - Avoiding grass and water (4 pts)
  - Gender recognition (2 pts)
  - Auton. space exploration (2 pts)
  - Dialogue with ASR (2 pts)
  - Pointing the camera (1pt)
  - Creating a bird catalogue (1pt)

- Performance evaluation
  - Navigation (1 pt)
  - Reasoning (1 pt)
  - Visualisation (1 pt)
  - Robustness (2 pts)
  - Relative speed (2 pts)
  - Overall impression (3 pts)

- Points:
  - Must do: 13
  - Should do: 12
  - Performance: 10
  - Total: 35

Task 2 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) in 3D
  - to learn and recognize colours
  - to detect and recognise birds
  - to classify faces according to gender
  - 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**