KIT-Department of Informatics
Prof. Dr.-Ing. Tamim Asfour

Exam Question Sheet

Robotics III - Sensors and Perception in Robotics

February 9, 2023, 17:30 – 18:30

- Please fill in your name and matriculation number clearly legible in the header of each answer sheet and the cover sheet.
- Exercise sheets will not be handed in. Therefore, enter your answers only in the areas of the answer sheets provided for each question. Answers on sheets submitted separately will not be graded.
- Apart from writing utensils, no other aids are permitted during the exam. Please use a permanent pen with black or blue ink. Answers written either with a pencil, with red or with green ink will not be graded. Attempts to deceive by using inadmissible resources will lead to exclusion from the exam and result in the grade „failed“.
- Unless otherwise stated in the question, please enter only the final results in the answer sheets. You can use the back sides of the question sheets as concept paper. Additional concept paper can also be provided on request during the exam.
- Please keep answers or explanations brief. The space provided on the answer sheets for a question does not correlate with the length of a correct answer.
- **Answers can be given either in English or German.** You are allowed to switch the language between answers, but not within an answer.
- The total score is 45 points.

Good luck!

Exercise 1 *Internal Sensors* (9 points)

1. Given is an optical absolute encoder with three binary code rings on the encoder disc. The read head of the sensors is currently at the position 111 and will change next to position 000.

   (a) Which erroneous transient states can occur during the change between the positions 111 and 000? 2 p.

   (b) How would you design an encoder disc in which such erroneous transient states do not occur? Explain why. 1 p.

   (c) What is the resolution of the encoder? 1 p.

2. Name and explain two measurement principles for the realization of a torque sensor that can be used in a robot joint. 2 p.

3. Which sensor modalities are included in an Inertial Measurement Unit (IMU)? How are the signals of the different modalities processed and fused to calculate the orientation? Which filters are used and why? 3 p.

Exercise 2 *External Sensors* (9 points)

1. What is the difference between RADAR and LiDAR? 1 p.

2. The output signal of a LiDAR sensor with *Amplitude Modulation* (LiDAR-AM) is shown in Figure 1.

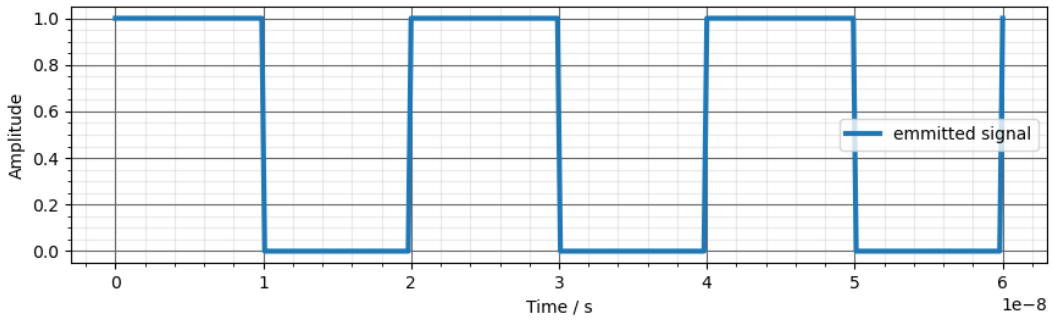

Figure 1: Emitted signal of the LiDAR-AM sensor.

   (a) The reflected signal is received with a phase shift of $\frac{\pi}{2}$. Draw the received signal with a dashed line in the solution sheet. 1 p.

   (b) Give the formula to calculate the distance d of the detected object. Give the result in meters (m). (Assume: $c = 100\,000\,\mathrm{km\,s^{-1}}$) 2 p.

   (c) Explain the term *Aliasing Effect* in the context of a LiDAR-based distance measurement. 1 p.

3. An RGB camera is connected to a computer using USB 3.0, which is capable of transmitting $5000\,\mathrm{Mbit\,s^{-1}}$. The image size is 1000×333 pixels with a 10 bit resolution per color. Calculate the maximum possible frame-rate for transmitting the raw camera images. 1 p.

4. Active Depth Cameras

   (a) Name three different projection schemes for *Structured Light Cameras*. 1 p.

   (b) Explain the term *Pattern Decoding* in the context of active depth cameras. 1 p.

   (c) Name two potential error sources for *Time-of-Flight* (ToF) systems. 1 p.

Exercise 3 *Feature Extraction* (9 points)

1. To find points of interest in an image, both the *Moravec Operator D* and the *Image Structure Tensor M* can be used. To evaluate which methods is better for object detection, you compare both methods using the example image I

$$I = \begin{pmatrix} I(0,0) & \dots & I(0,4) \\ \vdots & \ddots & \vdots \\ I(4,0) & \dots & I(4,4) \end{pmatrix} = \begin{pmatrix} 0 & 1 & 2 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 1 & 0 \end{pmatrix}$$

   (a) Write the equation for the *Moravec Operator* $D(u, v, s, t)$ and calculate D for $(s, t) \in \{(1, 0), (0, 1), (1, 1), (-1, 1)\}$ using a 3×3 window at pixel $(2, 2)$ of the image I. [2 p.]

   (b) For pixel $I(2, 2)$, the image derivatives I_x and I_y are given as

$$I_x(2,2) = \begin{pmatrix} 0 & 0 & 0 \\ -3 & -4 & -3 \\ 0 & 0 & 0 \end{pmatrix}, \qquad I_y(2,2) = \begin{pmatrix} -6 & 0 & 6 \\ -4 & 0 & 4 \\ -4 & 0 & 4 \end{pmatrix}$$

   Write the general equation for the *Image Structure Tensor* $M(u, v)$ in terms of $I_x(u, v)$ and $I_y(u, v)$ and calculate $M(2, 2)$. [2 p.]

   (c) How would you calculate the image derivatives I_x and I_y? [1 p.]

2. Scale Invariant Feature Transform

   (a) What are the four steps of the *SIFT* algorithm? [2 p.]

   (b) Describe how *SIFT* features can be used for 6D pose estimation of known objects. [2 p.]

Exercise 4 *Scene Understanding* (9 points)

1. What is the purpose of a *Region Proposal Network* in object detection? 1 p.
2. What is the output of *semantic* segmentation and *instance* segmentation, respectively? 1 p.
3. Machine learning methods are widely used for tasks like image segmentation and object detection.
   (a) What is the advantage of using point clouds for segmentation instead of color images? Describe a case in which point clouds are more suitable. 1 p.
   (b) Why is it challenging to apply classical machine learning architectures designed for images on point clouds or scene graphs? 2 p.
   (c) How do *PointNet* and *Graph Neural Networks* overcome these challenges? 1 p.
4. Object Relations are an important modality used for manipulation planning in unstructured scenes.
   (a) What is the difference between spatial and support relations? 1 p.
   (b) What is the difference between static spatial vs dynamic spatial relations? 1 p.
   (c) Name a perception task that can benefit from dynamic spatial relations as input. 1 p.

Exercise 5 *Active Vision* (9 points)

1. How would you distinguish between classical computer vision, active vision and active perception? Fill the table in the answer sheet. 2.5 p.
2. In the lecture, we discussed a method for the segmentation of unknown objects through physical interaction with the scene, which is based on

   > Schiebener, D., Ude, A. and Asfour, T., *Physical Interaction for Segmentation of Unknown Textured and Non-textured Rigid Objects, IEEE International Conference on Robotics and Automation (ICRA)*, 2014

   (b) Which heuristics can be used to obtain initial object hypotheses? Which type of objects are they suited for? 3 p.
   (c) What are the steps of ICP algorithm used to estimate the transformation between two point clouds? 1.5 p.
   (d) How can the standard ICP algorithm be improved to e. g. overcome shape ambiguities? 1 p.
   (e) Name two advantages for pushing the object multiple times. 1 p.