Karlsruhe Institute of Technology

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

Exam Question Sheet

Robotics III - Sensors and Perception in Robotics

March 2, 2021, 08:00 – 09:00

- 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* (10 points)

1. Given an optical incremental quadrature encoder with two tracks and four slits in the encoder disk.
   (a) How many ticks (number of rising and falling edges on both tracks) are generated for a full turn of the encoder disk? Explain your answer. 2 p.
   (b) Explain how the signals of the two tracks can be used to determine the direction of rotation. 2 p.
2. Absolute encoders
   (a) What is the advantage of an absolute encoder compared to a relative encoder. 1 p.
   (b) Draw a two bit absolute encoder disk on the solution sheet using binary encoding. 2 p.
   (c) What is the advantage of gray code encoding for encoder disks compared to binary encoding? 2 p.
   (d) Name two sensing principles used for absolute encoders apart from optical sensing. 1 p.

Exercise 2 *Tactile and Visual Sensing* (9 points)

1. In the lecture, the multi-modal fingertip sensor system developed at the H^2T was presented. Which type of sensors can be used to detect slip between object and finger? Which physical property is measured by the sensor? 2 p.
2. In the lecture, we introduced the tactile sensors *FingerVision* and *GelSight*. Both sensors use a camera to observe a set of markers on a deformable membrane.
   (a) Describe how an exerted torque perpendicular to finger surface displaces the markers and how the torque magnitude is computed. 2 p.
   (b) How can normal forces and tangential forces be distinguished from marker displacement information? 1 p.
3. Was is a *taxel*? 1 p.
4. In the lecture different types of proximity sensors were presented. Name one type of sensors that can detect the following materials robustly and a second that is not suited well for this task. Explain the reason why the sensor is not suited well. 3 p.
   - White cloth
   - Acrylic glass
   - Another robot

Exercise 3 *Feature Extraction* (8 points)

1. Calculate the *Sum of Squared Differences* and the *Sum of Absolute Differences* of the images I_1 and I_2. Are the correlation functions robust to outliers and invariant to differences in brightness? 3 p.

$$I_1 = \begin{pmatrix} 1 & 5 & 8 \\ 2 & 6 & 3 \\ 4 & 1 & 7 \end{pmatrix}, \qquad I_2 = \begin{pmatrix} 3 & 2 & 6 \\ 2 & 9 & 0 \\ 1 & 1 & 3 \end{pmatrix}$$

2. In Fig. 1, the results of two different corner detection operators are displayed. Name the operator, which results in the respective detected features and justify your decision. 2 p.

Figure 1: Results of two corner detectors

3. The Harris Corner Detector approximates the image function $I(u+s, v+t)$ using a first-order Taylor series

$$I(u+s, v+t) = I(u,v) + \begin{pmatrix} I_x(u,v) & I_y(u,v) \end{pmatrix} \cdot \begin{pmatrix} s \\ t \end{pmatrix}$$

   (a) Name two operators that can be used to numerically calculate I_x and I_y. 1 p.

   (b) The following is a section of the image matrix I: 1 p.

$$I = \begin{pmatrix} - & - & - & - & - \\ - & i_{u-1,v-1} & i_{u-1,v} & i_{u-1,v+1} & - \\ - & i_{u,v-1} & i_{u,v} & i_{u,v+1} & - \\ - & i_{u+1,v-1} & i_{u+1,v} & i_{u+1,v+1} & - \\ - & - & - & - & - \end{pmatrix} = \begin{pmatrix} - & - & - & - & - \\ - & 4 & 5 & 4 & - \\ - & 3 & 1 & 1 & - \\ - & 6 & 8 & 2 & - \\ - & - & - & - & - \end{pmatrix}$$

   Calculate $I_x(u,v)$ using the kernel

$$K_{I_x} = \begin{pmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{pmatrix}$$

   Hint: The result is a scalar.

   (c) The Eigenvalues λ_1 and λ_2 of the *Image Structure Tensor* $M(u,v) \in R^{2\times 2}$ indicate different regions in an image. Fill out the table on the answer sheet by providing the region type for the different Eigenvalue combinations. 1 p.

Exercise 4 *Scene Understanding* (10 points)

1. Name two challenges occurring when learning object relations that are difficult to handle when using traditional machine learning methods (e. g. Multilayer Perceptrons), which can be handled well by Graph Networks (GN). 1 p.

2. Given a GN block which operates on scalar edge attributes $e_k \in \mathbb{R}$ $(k \in \mathbb{N})$. Let $\rho^{e\to u}(E')$ be the GN block's edge aggregation function, which takes as an input the updated edge attributes $E' = \{e'_1, \ldots, e'_m\}$. 3 p.

   For each of the following functions, state whether it is a valid aggregation function or not. If a function is not valid, explain why.

$$(1) \quad \rho_1^{e\to u}(E') = \rho_1^{e\to u}(\{e'_1, \ldots, e'_m\}) = \sum_{k=1}^{m} (e'_k)^2$$

$$(2) \quad \rho_2^{e\to u}(E') = \rho_2^{e\to u}(\{e'_1, \ldots, e'_m\}) = (e'_1) + (e'_2)^2 + \cdots + (e'_m)^m$$

$$(3) \quad \rho_3^{e\to u}(E') = \rho_3^{e\to u}(\{e'_1, \ldots, e'_m\}) = \begin{pmatrix} \min\{e'_1, \ldots, e'_m\} \\ \max\{e'_1, \ldots, e'_m\} \end{pmatrix}$$

3. Given the graph G in Figure 2 with the node attributes $V = \{v_1, v_2, v_3\}$, the edge attributes $E = \{e_1, e_2, e_3\}$ and the global attribute $\mathbf{u}$, where 6 p.

$$v_1, v_2, v_3 \in \mathbb{R}, \qquad e_1, e_2, e_3 \in \mathbb{R}, \qquad \mathbf{u} = \begin{pmatrix} u_x \\ u_y \end{pmatrix} \in \mathbb{R}^2.$$

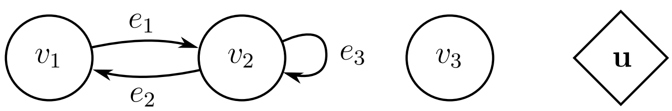

Figure 2: Graph G.

   The graph is passed to a Graph Network (GN) Block to update its attributes. The GN Block's update functions are defined as follows:

$$\text{Edge update: } e'_k = \phi^e\left(e_k, v_{s_k}, v_{r_k}, \begin{pmatrix} u_x \\ u_y \end{pmatrix}\right) := e_k + v_{s_k} \cdot u_x$$

$$\text{Vertex update: } v'_i = \phi^v\left(v_i, \begin{pmatrix} u_x \\ u_y \end{pmatrix}, \bar{e}'_i\right) := v_i - \bar{e}'_i + u_y$$

$$\text{Global update: } \mathbf{u}' = \phi^u\left(\begin{pmatrix} u_x \\ u_y \end{pmatrix}, \bar{v}', \bar{e}'\right) := \begin{pmatrix} u_x + \bar{e}' \\ u_y + \bar{v}' \end{pmatrix}$$

The GN Block's aggregation functions yield the *minimum* of their inputs (or 0 if the input set is empty). The graph's initial attribute values are:

$$v_1 = 3, \quad v_2 = 2, \quad v_3 = -1, \quad e_1 = 3, \quad e_2 = 0, \quad e_3 = -3, \quad \begin{pmatrix} u_x \\ u_y \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \end{pmatrix}$$

Calculate the updated attribute values and aggregation results.

Hint: Formally, all aggregation functions $\rho^{e \to v}(E'_i)$, $\rho^{v \to u}(V')$ and $\rho^{e \to u}(E')$ have the form:

$$\bar{x} = \rho(X') = \rho(\{x'_1, \ldots, x'_m\}) := \begin{cases} \min\{x'_1, \ldots, x'_m\}, & X' \neq \emptyset \\ 0, & X' = \emptyset \end{cases}$$

Exercise 5 *Active Vision* (8 points)

1. Which region of the human eye has the highest visual acuity. 1 p.
2. In the lecture, we introduced a transsaccadic memory model of visual search.
   (a) Name the layers in the transsaccadic memory introduced in the lecture. 1 p.
   (b) Insert the labels *Environmental Model*, *Gaze Selection* and *Saliency Calculation* in the correct box of the task-informed shared attention model given on the answer sheet. 1 p.
3. Given a fixation point x_{FP} for the current gaze direction, describe the basic steps of a gaze stabilization method that keeps x_{FP} centered and compensates for the motion induced by the robot's own movements. 2 p.
4. A robot observes a dynamic scene, where objects are moving in front of the robot. Give the control law of the gaze stabilization reflex you would implement. Which variable would you change if the implementation overcompensates, i.e., the generated eye velocities are too high? 3 p.