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

# Exam Solution Sheet

Robotics III - Sensors and Perception in Robotics

Febuary 15, 2022, 08:00 – 09:00

| Family name: | Given name: | Matriculation number: |
|---|---|---|
| Bond | James | 007 |

| | |
|---|---|
| Exercise 1 | 10 out of 10 points |
| Exercise 2 | 9 out of 9 points |
| Exercise 3 | 8 out of 8 points |
| Exercise 4 | 10 out of 10 points |
| Exercise 5 | 8 out of 8 points |

| | |
|---|---|
| **Total:** | 45 out of 45 points |

| | | |
|---|---|---|
| | **Grade:** | **1,0** |

# **Exercise 1** *Internal Sensors* (10 points)

1. Bandwidth:
   Range between the lowest and highest cutoff frequencies (slowest and fastest change in the input signal that can be correctly measured by the sensor).

   Response time: Time delay from change in input to change in output.

2. Optical Encoders:
   (a) Working principle:
   Light source shines/reflects on a detector. A partially transparent/reflective encoder disc periodically interrupts the light beam. The interruptions are detected and summed up to form the position values.
   (b) Direction detection:
   A second, phase-shifted code track enables detection of the direction of rotation (Quadrature Encoder).

3. Piezo-resistive effect:
   (a) Explanation: Materials change their electrical resistance when deformed.
   (b) Strain gauges: Strain gauges turn micro-deformations into a change of their electrical resistance. The effect is amplified by a series connection of many windings.
   (c) Piezo-resistive accelerometer: Seismic proof mass on deformable sensing elements. Deformation of the sensing elements is measured through piezo-resistive effect.

4. Accelerometer:
   Detect the effects of acceleration on a seismic proof mass. Most common: deflection measurement on a spring-mass system.

5. Filter:
   Accelerations that are not due to gravity cause (short term) errors in the accelerometer based orientation estimation. Angular rate gyroscopes measure only relative orientation w.r.t. initial orientation. Long-time drift caused by the numerical integration cannot be avoided. Low-pass filter on the accelerometer estimates and high-pass filter on the gyroscope before fusion of the signals minimizes sensor induced errors.

6. Yaw angle:
   Drift about the vertical axis (Heading) can not be compensated for as gravity does not provide any information there. A magnetometer provides a drift-free reference (magnetic north) for the rotation around the vertical axis (yaw)

# **Exercise 2** *External Sensors* (9 points)

1. Proximity Sensors:
   Capacitive, optical, inductive, acoustic

2. Difference:
   A proximity sensor provides a binary signal based on a threshold whether an object is present or not, a distance sensor outputs a distance value

3. Sensor for transparent objects:
   Acoustic proximity senor or capacitive proximity sensor

4. LiDAR Sensor

   (a) Reason for $\Delta f_D$: The frequency shift is caused by the Doppler effect: The movement of the obstacle causes the received signal frequency to be lower than the emitted frequency.

   (b) Object Distance: The object distance $D$ to the sensor can be calculated with the formula $D = \frac{\Delta t \cdot c}{2}$

5. LiDAR maps:
   The sensor is swept vertically and horizontally over the scene

6. Depth image:
   An infrared pattern is projected onto the scene. The deformed pattern is captured by an infrared camera. Depth values are calculated from the deformed pattern compared to the initial pattern.

7. Description of the case:
   As structured light cameras rely on the deformation of the projected pattern (active camera), they can also measure the distance to featureless surfaces. Therefore, when the scene does not contain optical features, e. g. , on a mono-colored surface, a stereo camera can not estimate the depth, a structured light camera can obtain depth values.

## Exercise 3 *Tactile Sensing* (8 points)

1. Tactile Perception: Sensation arising from stimulus to the skin (i. e. temperature, pressure, vibration, slip, pain) 2 p.
   Proprioception: Estimation of limb position/location, motion, force through end organs located in muscles, tendons and joints
   Haptics = Tactile + Proprioception

2. Weiss Robotics Tactile Sensor: Sensor changes resistance when force is applied. Resistance is high when force is low 1 p.

3. FingerVision & GelSight: Transparent layer with known marker positions. Cameras record displacement of marks due to contact forces. The displacement can be used to estimate force and torque.

4. Sensorized Robotic Fingers: 2 p.

| | |
|---|---|
| Joint angle sensors | Finger positions in underactuated hands can not be inferred by motor relative encoders |
| Distance sensing | Sense distance to the object when approaching it |
| Accelerometer | Used to detect vibrations induced by slip |
| Normal and shear force sensors | Needed to control the grasping force for fragile objects |

5. Contact Estimation : Contact sensing is established by adding a static force sensing modality in the form of a MEMS barometer. 1 p.

# Exercise 4 *Scene Understanding* (10 points)

1. Levels: 2 p.

   - Images (color, depth, point clouds)
   - Annotated images (classification, bounding boxes, pixel-wise labeling, segmentation)
   - Object instances (object instance detection, object localization, 6D pose estimation)
   - Object relations (spatial relations, temporal relations, support relations)

2. Methods: 2 p.

   - Sliding window: Move a window over the image systematically / on a regular grid / exhaustively and apply $M$ to each sub-image to detect a person in this window.
   - Region proposal: Use another method (e.g. neural network) to predict candidate regions and apply $M$ to each candidate region.

3. Challenge: 1 p.

   - In general 3D point clouds, there is no natural order or grid-like structure between the elements, i. e. points, while neural networks for images rely on a grid-like structure of the pixels.
   - A point cloud can have an arbitrary and varying number of points, while images have a constant number of pixels.
   - The density of points in space can vary greatly, while pixels are distributed evenly over the image space.

4. Difference: 1 p.

   For pixel-wise segmentation, each pixel must be assigned a class (e. g. vehicle, person, street, building, . . . ). Therefore, different instances of the same class (e. g. two cars) are assigned the same label. For instance segmentation, different instances of the same class must be assigned different labels.

5. Support Graph

   (a) Method: 2 p.

   - Topological sorting of the nodes
   - Find sequence $(o_1, ..., o_n)$ so that $\forall i, j$ with $i < j$, $(o_i, o_j) \notin E$
   - Find sequence of objects so there is no support relation / edge from an object removed earlier to an object removed later.
   - Find a leaf node / node with out-degree 0 / node without outgoing edges (must exist because graph is acyclic), remove it from the graph and add it to the sequence. Repeat until all nodes are removed.

   (b) Change: 1 p.

   With cycles, there is no such sequence/sorting/ordering. The robot needs to remove multiple objects at a time or manually secure an object while removing another.

6. Difference: 1 p.

   Static relations only depend on the relative position of objects at one point in time , while dynamic relations also depend on the previous / initial relative position of both objects .

# Exercise 5 *Robot Vision* (8 points)

1. Foveal vision in robotics: Two cameras per eyes. One with a higher focal length for foveal vision and the other with a lower focal length for peripheral vision. Other solution: log-polar cameras.

2. Feature matching descriptors
   (a) SIFT
   (b) ORB
   (c) SURF
   (d) MSER

3. Maximally Stable Extremal Regions are defined solely by the intensities of an image. The idea is to find regions that remain consistent over a wide range of intensity thresholds. Then, take the most stable version of a consistent region.

4. The active visual search problem is NP-complete.

5. Reafference: