Final Examination

Applied Computer Vision

Summer Term 2022

Technische Universität Berlin | Computer Vision & Remote Sensing

Name: . . . . **Student ID:** . . . .

Altogether ***33 points*** can be obtained. You have ***90 minutes*** to complete this exam.

- Please fill in your name and student ID on all sheets.
- No auxiliary resources allowed**.**
- Use a blue or black ballpoint pen - no pencil. Do not use any kind of correcting fluid or tape.
- A short and accurate style as well as a clear handwriting should be intended.
- Pay attention to a clear and comprehensible preparation of sketches.

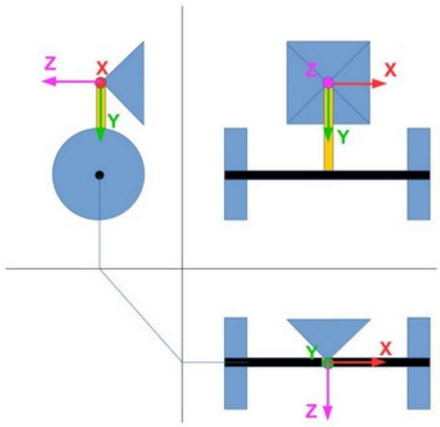
1. Predicting Optical Flow

14P

For the robot car equipped with a camera as shown in the Figure below you want to predict optical flow in a simulation computation. The Figure is drawn to scale. The diameter of a wheel is 5 cm.

Left and right robot wheels are driven by two motors whose velocities can be set to v_l and v_r in rotations per second (rps).

The camera provides quadratic images with a size of 1000 x 1000 pixels. The projection center of the camera is located 500 pixels in front of the image plane.



a) Give the center points of the robot wheels $\mathbf{P}_l$ and $\mathbf{P}_r$ in the camera's 3D coordinate system. As shown in the Figure, the camera's X axis is parallel to the wheels' axis. The Z axis is perpendicular to the wheels' axis and horizontal in world coordinate system and points forwards. And the projection center is located above the middle of the wheel's axis. 3P

b) If (at time t = 0) you set the left wheel's speed to v_l = 1 rps and the right wheel's speed to v_r = 1 rps, what is the velocity (i.e. speed) of the camera's projection center $v_\mathbf{P}(t)$ in cm/s as a function of time t? In addition to this, please also give the velocity vector $\mathbf{V_P}(t) = (v_x, v_y, v_z)^T$ at time t = 3 sec in the 3D coordinate system of the camera at time t = 0 (Hint: the answer to the latter question is almost trivial considering the situation as described.). 2P

c) In case of robot motion according to b), what optical flow $o(\mathbf{x}) = (v_x, v_y)^T$ in pixels will be observed at the center/mid point of the image $\mathbf{x}=(0, 0)^T$? (Optical flow is measured between two subsequent video frames acquired with a frame 1P frequency of 30 Hz.)

d) If you set the left wheel's speed to v_l = 1 rps and the right wheel's speed to v_r = -1 rps, what is the optical flow o(**x**) = $(v_x, v_y)^T$ in pixels at the center/mid point of the image **x**=$(0, 0)^T$? (Optical flow is measured between two subsequent video frames acquired with a frame frequency of 30 Hz.) 3P

e) One way to describe the orientation of a camera in a spatially fixed 3D world coordinate system is the projection matrix **P**. When the robot car moves, **P**(*t*) changes as a function of time *t*. Assuming constant speed settings for the two robot wheels, what are the parameters **P**(*t*) depends on? (Hint: Do not derive the complete formula, but only name all the parameters – geometric and others – **P**(*t*) depends on.) 3P

f) What changes compared to question **e)** if speed settings $v_l(t)$ and $v_r(t)$ are not constant any more, but change as a function of time *t*? How would you incorporate the change into your numerical computation of P(t)? 2P

2. PID Controller

5P

We are using a PID Controller to steer our robot along a given a line.
We measure the deviation from the desired position as the error **e(t)**.

Given are the following measurements: *(e(0), e(1), e(2), e(3), e(4)) = (-5,-4,0,2,6)*
The time between measurements is Δt = 1.0 s.

1P

a) Let K_p*=1.5* be the proportional parameter. What is the proportional term's contribution to the update computed by the PID controller at *t=4*?

1P

b) Let K_I*=0.5* be the integral parameter. What is the integral term's contribution to the update computed by the PID controller at *t=4*?

1P

c) Let K_D*=0.2* be the derivative parameter. What is the derivative term's contribution to the update computed by the PID controller at *t=4*?

1P

d) What is the control signal *u(t)* at *t=4*?

1P

e) Due to the image resolution the calculated error has a maximum range of $e(t) \in [-100,+100]$. From empirical experiments the range for the integral term is $[-200,+200]$ and the range for the derivatives is $[-50,+50]$. However, the robot controller only allows control signals in the range $u(t) \in [-5,+5]$. What are suitable parameters K_p, K_I, K_D , such that the control signal is scaled to the desired range without relying on any thresholding?

Explain your reasoning!

3. Visual Odometry

10P

You have a video from a camera mounted to a robot. You don't know its odometry, i.e. you don't know its control signals and have no motion sensors. You are continuously computing optical flow images between two consecutive frames, which provide for each frame ***t*** a collection of feature points $\{(x_i^{(t)}, y_i^{(t)})\}_{i=1}^N$ and corresponding offset vectors $\{(dx_i^{(t)}, dy_i^{(t)})\}_{i=1}^N$

2P

a) Before mounting the camera we estimated the intrinsic parameters of the camera.

$$K = \begin{bmatrix} \phi_x & \gamma & \delta_x \\ 0 & \phi_y & \delta_y \\ 0 & 0 & 1 \end{bmatrix}$$

Name and describe each parameter.

2P

b) Illustrate the epipolar geometry between two images.

c) Describe the relationship between two points that fulfill the epipolar constraint mathematically. **2P**

d) Given the optical flow and intrinsic parameters, explain how to estimate the movement of the robot between two frames. Describe all geometric/algebraic parameters that need to be computed. **4P**

4. Deep Learning

4P

You are building a computer vision system for a self-driving car using Deep Learning. The goal is to detect other cars, pedestrians and various traffic signs. Additionally, the model should be able navigate autonomously.

2P

a) Explain how you would implement a system to detect objects. Describe the creation of the dataset, the model and its training. Describe as detailed as possible what the model gets as input and what it outputs.

Briefly explain how the model could be used for navigating the car.

b) We decide to steer the car directly without relying on any object detection. The model outputs a steering command given visual information. **2P**

Explain how you would implement a system to drive autonomously. Describe the creation of the dataset, the model and its training. Describe as detailed as possible what the model gets as input and what it outputs.

Briefly discuss how the model could be used for navigating the car. What advantages and disadvantages exist to the method in **a)** ?