

**Affix the authentication label over this field.**

Surname: .....

First name: .....

Matr.-No.: .....

**Institut für Mess- und Regelungstechnik  
mit Maschinenlaboratorium  
Karlsruher Institut für Technologie (KIT)  
Prof. Dr.-Ing. C. Stiller**

**Exam in “Machine Vision”**

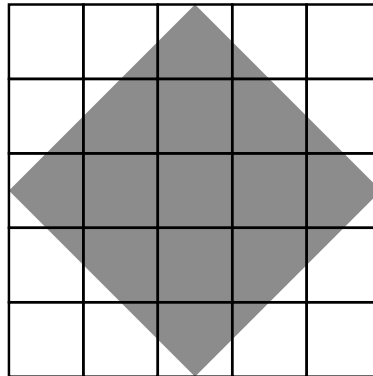
Date of exam:      September 14, 2023  
Time of exam:      8:00-9:00

---

Question 1

(6 points)

The picture below shows a 5-by-5 pixel grid with a gray diamond shaped area. Design a filter mask that calculates the average gray value over the diamond shaped area.





---

Question 2

(4 points)

Design a potential function for a random field based segmentation that fosters that a pixel shares the same label as its neighbors if all eight neighbors share the same label. The potential function should not have any effect if the eight neighbors do not agree on the same label. Denote with  $l_0$  the label of the pixel of concern and with  $l_1, \dots, l_8$  the label of the eight neighboring pixels. Furthermore,  $c_0, \dots, c_8$  denote the colors of the pixels.



---

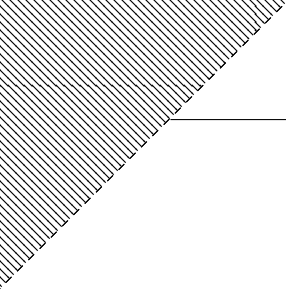
Question 3

(4 points)

Assume that the loss term  $g(w)$  of an artificial neural network has the form

$$g(w) = w^3 - 10w + 20$$

Perform two steps of gradient descent with learning rate  $\varepsilon = \frac{1}{5}$  starting from the initial weight  $w_0 = 0$ . For each step provide the value of  $w$  and the value of  $\frac{\partial g}{\partial w}$ .



---

Question 4

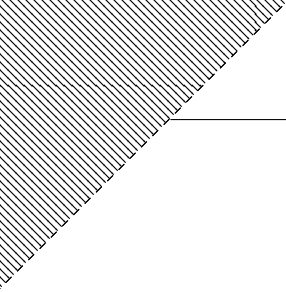
(3+3 points)

Assume a single perceptron with ReLU activation and two inputs  $x_1$  and  $x_2$ . We consider as input only values in the interval  $-1 \leq x_1, x_2 \leq 1$ . Other inputs are not considered. Which of the following functions can be implemented by the perceptron? Explain your answer briefly.

(a)  $f(x_1, x_2) = x_1 + 2x_2 + 3$

(b)  $g(x_1, x_2) = \begin{cases} x_1 + x_2 & \text{if } x_1 > 0 \text{ and } x_2 > 0 \\ 0 & \text{otherwise} \end{cases}$





---

Question 5

(8 points)

Implement a Python function *make\_binary\_histogram* that calculates the binary histogram for local binary patterns from an image. The function should take as input an eight-bit encoded graylevel image and should yield a 256-dimensional vector that implements the histogram. It should consider all pixels in the image except of the boundary pixels.

You might make use of a function *make\_decimal* that takes as input a binary vector (i.e. a vector composed out of entries 0 and 1) that is interpreted as binary number.

*make\_decimal* returns this number. Example: `make_decimal([ 0,0,1,1,0,0,1,1 ])` returns 51.

```
import numpy as np

def make_decimal(pattern: np.ndarray) -> int:
    v = 2*np.array([7,6,5,4,3,2,1,0])
    return np.sum(pattern * v)

# @student: implement function make_binary_histogram
```



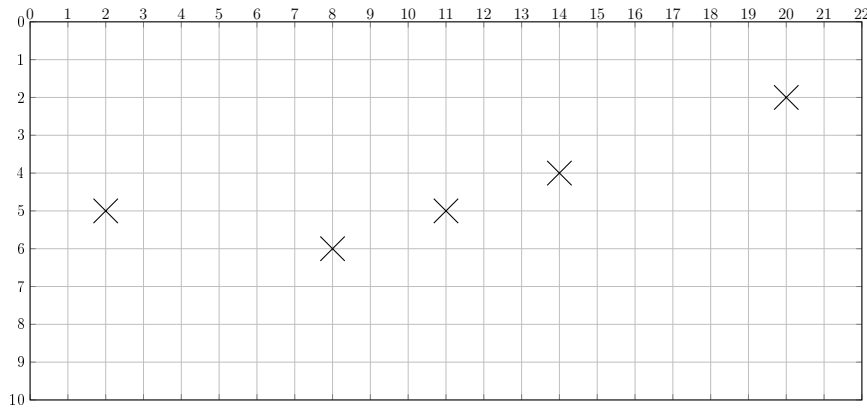
---

Question 6

(5+3 points)

A linefitting must be applied to five sample points which are displayed below. To reduce the impact of outliers the RANSAC algorithm is used. All possible combinations of sample points are used to check the number of outliers in the tolerance band. The threshold  $\theta$  can be either set to 0.5 or to 2 pixels.

- (a) Which combination of two points and a threshold  $\theta$  returns the best result?
- (b) Assume that all sample point that are not outliers are perceived with very low noise. What threshold  $\theta$  must be chosen and why? Which sample points provide the best result?





---

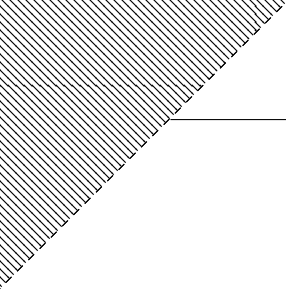
Question 7

(3+2+3 points)

A scene is illuminated by bluish light. To achieve color constancy, white balance is applied. A white reference surface is used to calculate average values for RGB.

- (a) Are the correction factors  $c_R$  and  $c_B$  less than, greater than or equal to 1?
- (b) How does the white balance effect the image brightness.
- (c) What happens if the lighting changes to greenish light and white balance is performed again?

Justify your answers briefly.



---

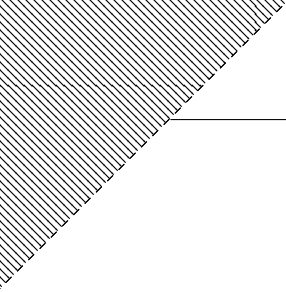
Question 8

(3+4+1 points)

Assume a camera with thin lens. The focal length of the lens is  $f_{lens} = 10$ . The focal length of the camera is  $f_{camera} = 11$ . The diameter of the aperture is  $D = 20$ .

- (a) Calculate the optimal object distance  $z_0$  at which points create a perfectly sharp image in the camera
- (b) Based on your result from part (a) calculate how the diameter of the unsharpness circle in the image plane  $\varepsilon$  depends on the object distance  $z$  for distances  $z \geq z_0$ . The unsharpness circle is the area on the image plane which is illuminated by a point light source at distance  $z$ .
- (c) Based on your result from part (b) how large can  $\varepsilon$  become at most for  $z \geq z_0$ ?





---

**Question 9****(4+4 points)**

Assume a binary classification task and an ensemble classifier with 20 ensemble members that was trained for that classification task. We apply the ensemble to a test set of 12 examples (named as A-L) and obtain the classification results as follows

test pattern	true class label	number of positive votes	number of negative votes
A	positive	17	3
B	positive	15	5
C	positive	9	11
D	positive	14	6
E	positive	20	0
F	positive	7	13
G	negative	15	5
H	negative	7	13
I	negative	2	18
J	negative	0	20
K	negative	1	19
L	negative	8	12

The ensemble classifier is comparing the number of positive votes minus the number of negative votes against a threshold  $\delta$ .

- (a) Calculate the precision and recall of the ensemble classifier for threshold  $\delta = 0$
- (b) Which values of  $\delta$  maximize the recall? What is the highest precision that we can achieve when we maximize the recall?

