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

Exam in "Machine Vision"

Date of exam: February 14, 2020
Time of exam: 16:00–17:00

(5 points)

The images below show a barcode which encodes the term 'machine vision' as well as a photo of the barcode. Unfortunately the barcode in the photo cannot be detected by a barcode reader. Which effect with respect to signal theory can be observed in this case? Explain why it is not possible to read the barcode from the photo.

How could this effect be avoided? Name two possible ways to take a better photo.

Question 2 (8 points)

Implement the non-maxima suppression approach of the Canny edge detector as a MATLAB function. The function should take the following arguments

- two real-valued image matrices `dgdu, dgdv` that contain the partial derivatives in horizontal and vertical direction for all pixels of the image

- a pixel position `u, v` for which the non-maxima suppression should be applied. You can assume that the function is not applied to pixels at the boundary of the image.

The function should return either `0` if the pixel should be suppressed by the non-maxima suppression or `1` if the pixel should not be suppressed.

You might use a function called `relative` that takes a 2d vector `(x,y)` and returns a 2d vector from the set $\{(-1,-1),(-1,0),(-1,1),(0,1),(1,1),(1,0),(1,-1),(0,-1)\}$ that points to the neighboring pixel with the most similar direction as `(x,y)`, e.g. `relative(2.3,0.2)` yields $(1,0)$, `relative(-1.3,1.4)` yields $(-1,1)$, `relative(1,-20)` yields $(0,-1)$.

Add your solution below the following function header

```
function [ q ] = non_maxima_suppression (dgdu, dgdv, u, v)
```

Consider the four points $\vec{x}_i$ given below. With RANSAC three lines l_1, l_2, and l_3 have been generated as possible candidate lines. Calculate the distances from each point to each candidate line and evaluate the number of outliers for each of the three candidates. Use an outlier threshold of $\theta = 1$. Which line fits best?
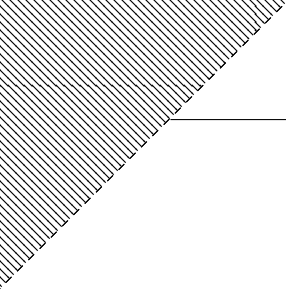
Remark: Note that the line equations are not normalized.

$$\vec{x}_1 = \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \quad \vec{x}_2 = \begin{pmatrix} 0 \\ -0.5 \end{pmatrix}, \quad \vec{x}_3 = \begin{pmatrix} 7.5 \\ 0 \end{pmatrix}, \quad \vec{x}_4 = \begin{pmatrix} -3 \\ 0.5 \end{pmatrix}$$

$$l_1 : \left\langle \begin{pmatrix} -1 \\ 4 \end{pmatrix}, \vec{x} \right\rangle + 2 = 0$$

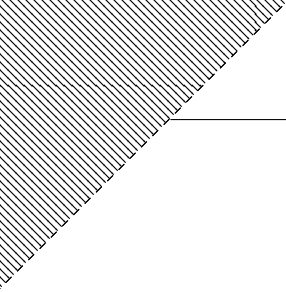$$l_2 : \left\langle \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \vec{x} \right\rangle + 1.5 = 0$$

$$l_3 : \left\langle \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \vec{x} \right\rangle = 0$$

Question 4 (4 points)

Explain why surfaces of colored objects appear differently if (a) they are seen outside in bright sunlight, and if (b) they are seen inside of a building without windows that is illuminated only by white fluorescent tubes.

Question 5 (5+2 points)

Develop a level set based segmentation method that mimics the region growing algorithm on graylevel images. Pixels should be added to the foreground segment if their gray value does not deviate more from the average gray value of the foreground segment than a given threshold $\theta > 0$.

(a) Provide an evolution dynamics for the boundary pixels, i.e. specify $\frac{\partial \vec{x}}{\partial t}$. You might use the variables g (gray value of pixel at position $\vec{x}$), $\bar{g}$ (average gray value in foreground segment), θ (threshold), ϕ (level set function), and $\varepsilon > 0$ (step width of evolution step).

(b) Provide the initial signed distance function if we want to start the evolution with a seed pixel at position $\vec{x}_0$.

Question 6 (6+1 points)

Assume a camera with thin lens. The lens has a focal length of $f_{lens} = 5$ mm. The diameter of the aperture is $D = 10$ mm and the tolerance diameter for blur on the image plane is $\epsilon = 0.01$ mm. Assume we adjust the camera in such a way that objects at a distance of 5005 mm are mapped perfectly sharp to the image plane.

(a) Calculate the hyperfocal distance d_h of the camera, the nearest object distance g_{near} at which objects are mapped to an area on the image plane that has a diameter of at most ϵ, and the largest object distance g_{far} at which objects are mapped to an area on the image plane that has a diameter of at most ϵ.

(b) Does the depth of field increase or decrease if we increase the diameter of the aperture D? (No explanation necessary)

A classification problem is solved to distinguish between four different object classes A, B, C, and D. The patterns are two-dimensional. For this reason, four classifiers were trained using the one-versus-the-rest approach as follows:

classifier	class	linear classifier
c_1	A	$x_1 - 3 > 0$
c_2	B	$2x_2 - x_1 - 2 > 0$
c_3	C	$-x_2 - 2 > 0$
c_4	D	$-3x_2 - x_1 - 3 > 0$

Add the decision boundaries of the four classifiers to the pattern space below. Label each area of the pattern space with the class (A, B, C, and D) which is predicted by the ensemble classifier. If the ensemble classifier is undecided for an area, list all the classes for which the ensemble classifier has at least one vote. Mark areas without any vote with the letter 'N'.

Draw carefully! In case you want to revise your solution, cross out Figure 1 and use Figure 2 instead.
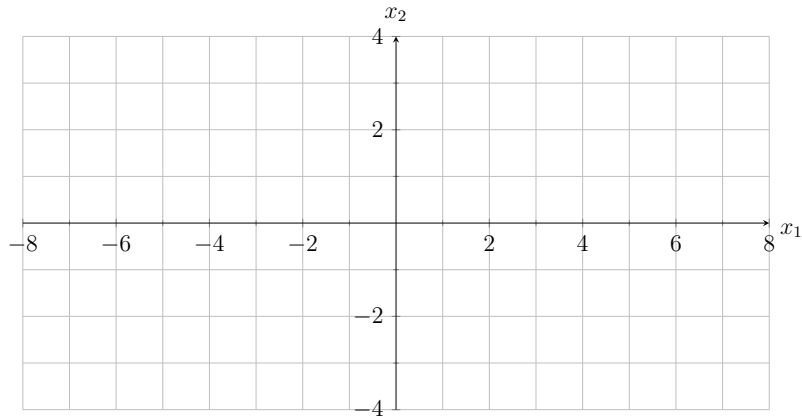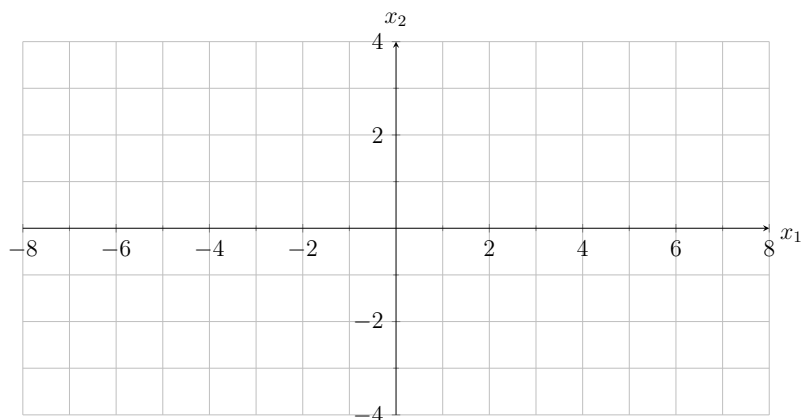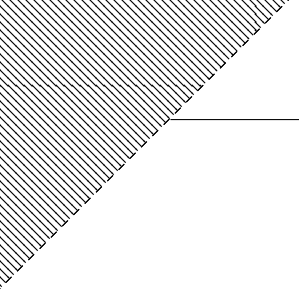
Figure 1: Enter your solution here.

Figure 2: Spare image just in case you need it (e.g. if you want to revise your solution).

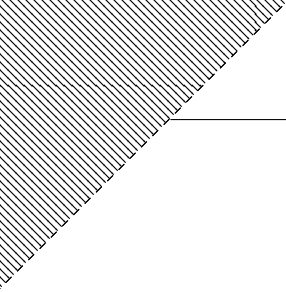Assume a convolutional neural network that is described in the table below.

layer	size (width×height×channels)	activation function	size of convolution kernel
1. input layer	$800 \times 600 \times 3$	-	-
2. convolutional layer	$800 \times 600 \times 48$	tanh	21×21
3. convolutional layer	$800 \times 600 \times 128$	tanh	11×11
4. convolutional layer	$800 \times 600 \times 256$	tanh	5×5
5. convolutional layer	$800 \times 600 \times 256$	tanh	1×1
6. fully connected layer	$100 \times 1 \times 1$	tanh	-
7. fully connected layer	$30 \times 1 \times 1$	softmax	-

(a) Calculate from which area of the input image the perceptrons at position $u = 400$, $v = 300$ of the fifth layer receive input. Provide this area as a rectangular box with coordinates of the left upper corner, width, and height. You may assume that all weights of the neural network are non-zero.

(b) Provide the area of the input image from which the first perceptron in the seventh layer of the network receives input. Justify your answer briefly.

Question 9   (2 + 3 points)

The plot below shows both the errors for the training and the validation dataset during the training of a neural network.

(a) Four different points in time t_i are highlighted in the plot. At which of those points in time does underfitting occur? At which of those points in time does overfitting occur?

(b) At which point in time should the training be stopped? Why will further training not be beneficial. What is the name of this technique?