DATE: 7. Dec. 2018

University of Stavanger

Faculty of Science and Technology

EXAM IN: ELE510 IMAGE PROCESSING with ROBOT VISION

DURATION: 4 hours, 09.oo - 13.oo

ALLOWED REMEDIES: Defined, simple calculator permitted.

THE SET OF EXERCISES CONSISTS OF 4 EXERCISES ON 8 PAGES

NOTES: Formulas are found on pages 9 - 10.

Exercise 1

(25%) This problem contains 10 questions that should have short concise answers.

a) In a digital camera the light intensity of the scene is transferred to a digital image. Give a short description of the two processes that is involved, *image sampling* and *image quantization*.

b) What is the purpose of a *Bayer filter*?

c) The point spread function of a certain filter can be written $h(i,j) = h_1(i)h_2(j)$. What is such a filter called?

d) Below is given three filter kernels. What is the purpose for each of these?

$$\mathbf{k}_1 = \begin{bmatrix} 0.075 & 0.124 & 0.075 \\ 0.124 & 0.2038 & 0.124 \\ 0.075 & 0.124 & 0.075 \end{bmatrix}, \qquad \mathbf{k}_2 = \begin{bmatrix} 0.17 & 0.67 & 0.17 \\ 0.67 & -3.36 & 0.67 \\ 0.17 & 0.67 & 0.17 \end{bmatrix} \quad \text{and} \tag{1}$$

$$\mathbf{k}_3 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ -0.075 & -0.124 & 0 & 0.124 & 0.075 \\ -0.124 & -0.204 & 0 & 0.204 & 0.124 \\ -0.075 & -0.124 & 0 & 0.124 & 0.075 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}. \tag{2}$$

e) Describe how filters are used to compute the *image gradient*.

f) The *Harris corner detector* uses the *Hessian matrix* (formula (8)). Explain how this matrix distinguish edges from corners.

g) What do we mean by the *image histogram*? Use the following equation in your explanation:

$$h(g_k) = n_k \qquad k \in \{0, 1, 2, \cdots, (G-1)\}. \tag{3}$$

h) In Figure 1 an image and its *power spectrum* is given. Explain the meaning of the lines in the power spectrum and how they are related to details in the image.

i) What is a *derivative of Gaussian filter*? Compute the filter kernel when the Gaussian kernel is

$$G_\sigma(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}}. \tag{4}$$

j) What is the following equation called? What is the use of this equation?

$$(\nabla I)^T \mathbf{v} + I_t = 0. \tag{5}$$

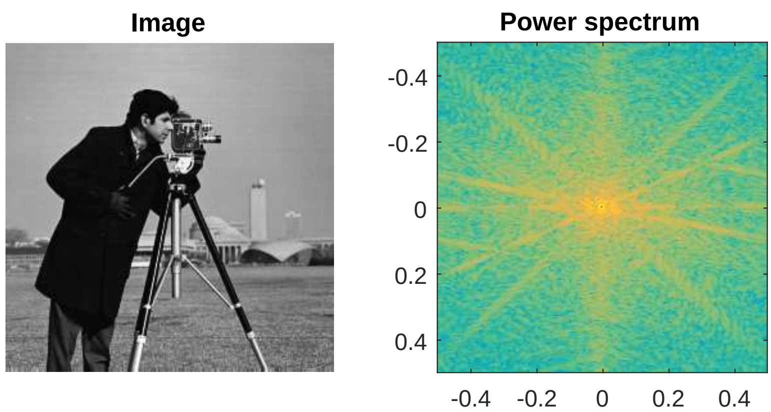

Figure 1: Left: image. Right: power spectrum

Exercise 2

(25%)

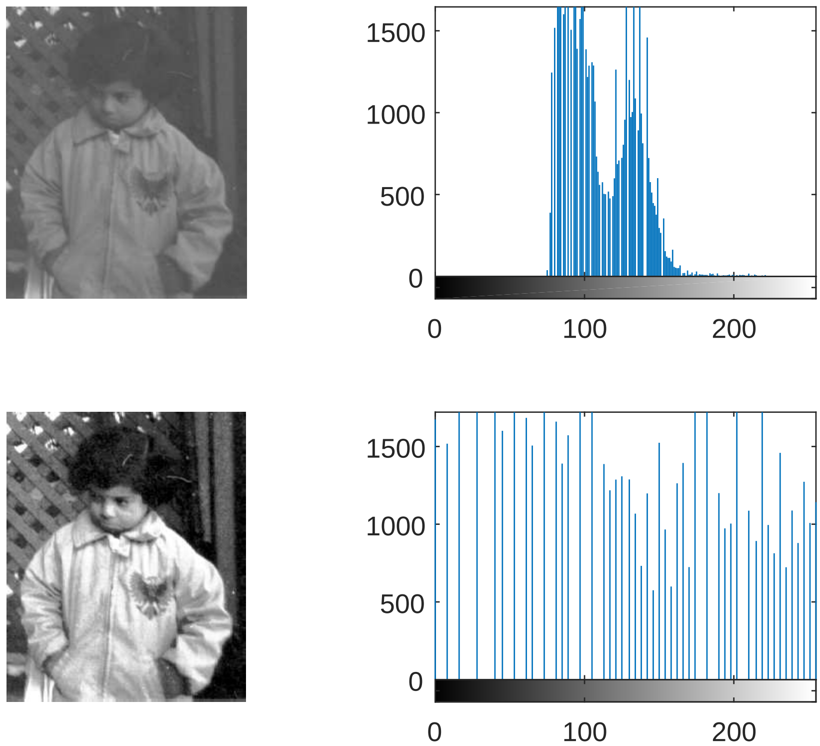

Figure 2: Upper; Left: image, Right: histogram. Lower: Left: output image and Right: histogram.

a) What is the purpose of *image enhancement*? Mention a few methods and give a short description.

b) In Figure 2, an image and its *histogram* is shown before and after enhancement. What is the method used here called, and how does it work?

c) In Figure 3, an image and the result from *edge detection* is shown. The *Canny edge detector* is used. Give a description of this edge detector. Write pseudo code.

d) What is *hysteresis thresholding*, its purpose? Explain how it is implemented?

Figure 3: Left: image, Right: binary edge image.

Exercise 3

(25%)

An indoor environment with a corridor and passages is shown in Figure 4. Positions and sizes can be found from the Figure (1.0 m grid size). A transporter robot, A, is moving through the corridor. The robot has a digital camera placed as shown in Figure 4. The camera can rotate in the X-Y-plane. The optical axis is defined with direction from the camera center towards the image plane. For the position shown in the Figure, the direction of view is therefore towards the opening (door) in the end of the corridor (i.e. negative values of Z_C). The world coordinates are defined with the Z axis vertical upwards with origin in the ground plane, in the lower left corner of the corridor. At time $t = 0$ the optical center of the camera is exactly as shown in Figure 4. The world X coordinate is parallel to the camera coordinate X_c. The homogeneous camera coordinates are $\mathbf{P}_c = [X_c \ Y_c \ Z_c \ 1]^T$. The image sensor has 3000×2000 pixels in an area $30\,\text{mm} \times 20\,\text{mm}$ (horizontal $\times$ vertical). Assume zero skew.

The field of view (FOV), given by angles in the $Z_c - X_c$ and $Z_c - Y_c$ plane is 53.13 degrees in the horizontal plane and 36.87 degrees in the vertical plane.

a) Compute the *focal length* for the camera.

b) Show that the internal calibration matrix $\mathcal{K}$ is

$$\mathcal{K} = \begin{bmatrix} 3000 & 0 & 1500 \\ 0 & 3000 & 1000 \\ 0 & 0 & 1 \end{bmatrix}. \tag{1}$$

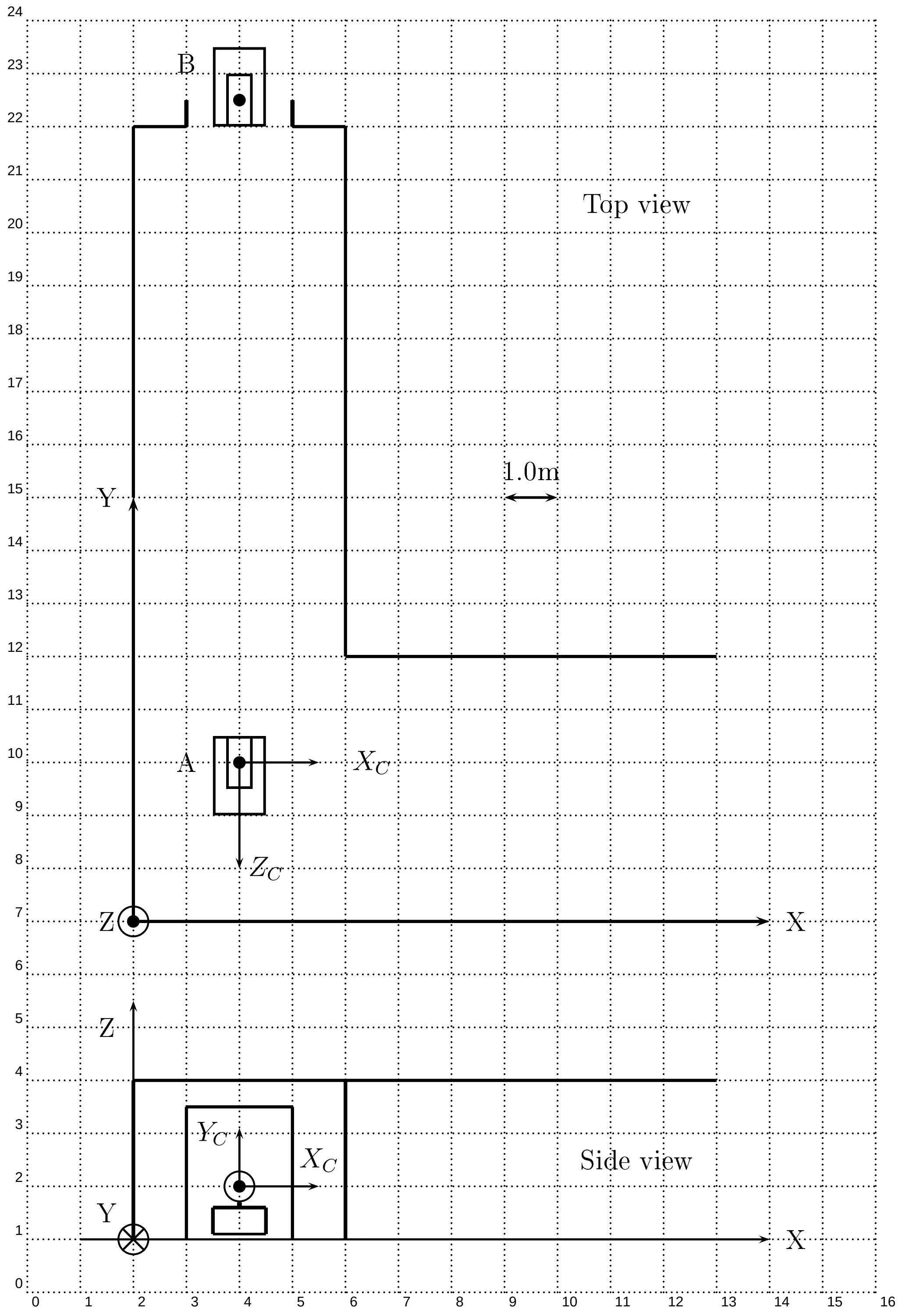

Figure 4: Passages with transportation robots.

c) Find the extrinsic parameters and show that the normalized ($m_{34} = 1$) camera matrix is given by

$$\mathcal{M} = \begin{bmatrix} 1000 & -500 & 0 & -500 \\ 0 & -333 & 1000 & 0 \\ 0 & -0.333 & 0 & 1 \end{bmatrix}. \tag{2}$$

An identical robot, B, is moving into the door, in the opposite direction of the first robot, at the end of the corridor. The position of this robot is as shown in Figure 4. The camera center for this robot is $\mathbf{P}_0^B = [2 \;\; 15.5 \;\; 1.0 \;\; 1]^T$ at $t = 0$.

d) Find the coordinates in pixels for the image of robot 2 in the camera of robot A. Use the corners of the rectangular side view and the camera center. Sketch the result.

e) If the camera on robot A rotates an angle θ around the Y_c axis, how will this influence the extrinsic camera parameters. Find an expression for the new rotation matrix and translation vector. See hint below.

Hint: Let the old camera coordinates in 3D be $\mathbf{P}_{c1}$ and the new camera coordinates in 3D $\mathbf{P}_{c2}$. Likewise let the world point in 3D be given by $\mathbf{P}_w$. Then

$$\mathbf{P}_{c2} = \mathbf{R}_2 \, \mathbf{P}_{c1} \tag{3}$$

and

$$\mathbf{P}_{c1} = \mathbf{R}_1 \, \mathbf{P}_w + \mathbf{t}_1, \tag{4}$$

where the rotation matrix $\mathbf{R}_1$ and the translation vector $\mathbf{t}_1$ is the result from subquestion c) above. $\mathbf{R}_2$ is the rotation matrix for the rotation of the camera an angle θ around the Y_c axis.

Exercise 4

(25%)

Consider the situation where robot A is stationary at the position in Figure 4. Robot B moves forward in a straight line towards robot A. On robot B there is two lights in the front with centers at $\mathbf{P}_1 = [1.5 \;\; 15.0 \;\; 0.5]^T$ m, right light, and the left light $\mathbf{P}_2 = [2.5 \;\; 15.0 \;\; 0.5]^T$ m, at $t = 0$. The speed of the robot is $V = 2$ m/s. The image frame rate is 40 frames per. second. As a help to the solution of the following questions, Figure 5 and Figure 6 are found below. We can find the position of the lights (in the image plane of camera A) as a function of time by using the equation

$$\lambda \mathbf{p}(k) = \mathcal{M} \mathbf{P}_w(k). \tag{1}$$

For the right light this gives the following result

$$x(k) = 1500 + \frac{500}{4 - \frac{0.05}{3}k}, \tag{2}$$

$$y(k) = 1000 + \frac{500}{4 - \frac{0.05}{3}k}, \tag{3}$$

a) Use equation (1) and find the scaling factor λ as a function of frame index k. First you have to find $\mathbf{P}_1(k)$.

b) Compute the optical flow for the right light point at $t = 4.0$ seconds.

c) What is the structure of the optical flow field for the total image in this case? What is the point in the center of Figure 5 called?

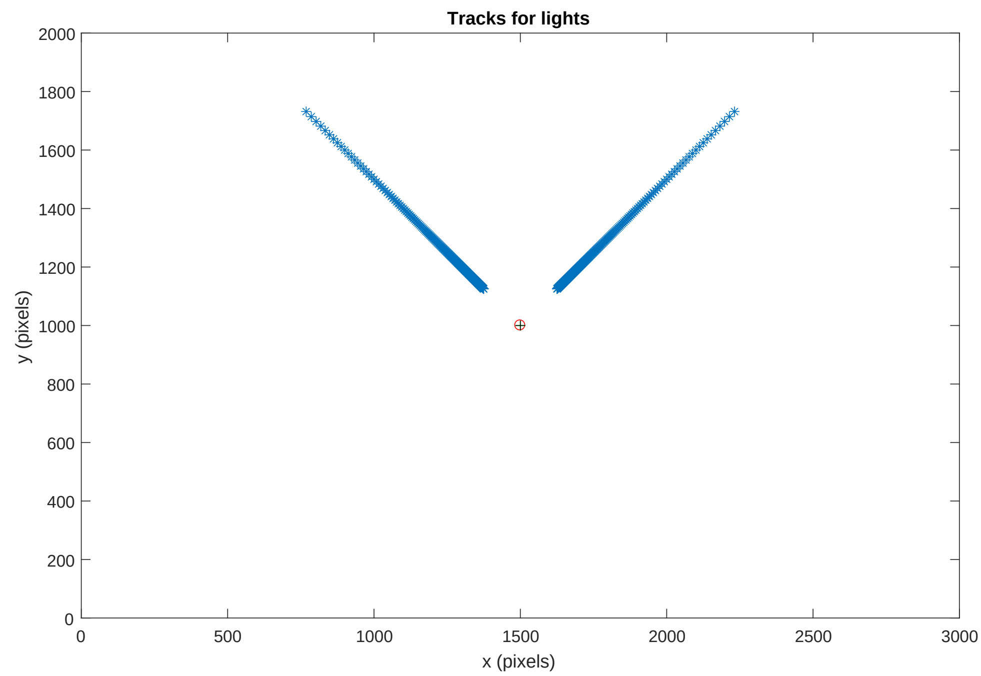

Figure 5: Tracks for the two lights from t=0 to t=5s.

d) How can we use the image information to find the time to collision?

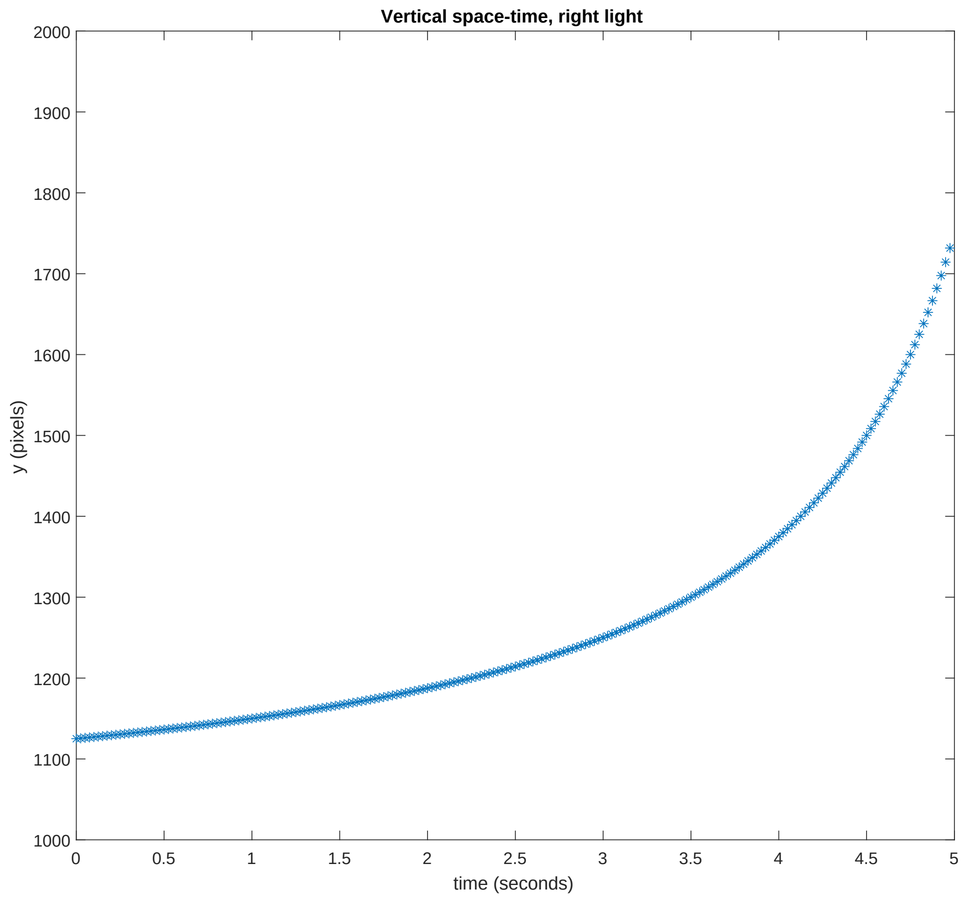

Figure 6: Space-time plot for vertical motion, y(t), from t=0 to t=5s.

Formulas

SVD decomposition

$$f = U\Lambda^{\frac{1}{2}}V^T \tag{4}$$

Discrete Fourier transform (DFT) and the invers DFT:

$$\hat{g}(m,n) = \frac{1}{MN}\sum_{k=0}^{M-1}\sum_{l=0}^{N-1} g(k,l)e^{-j2\pi[\frac{km}{M}+\frac{ln}{N}]} \tag{5}$$

$$g(k,l) = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} \hat{g}(m,n)e^{j2\pi[\frac{km}{M}+\frac{ln}{N}]} \tag{6}$$

The 2D convolution formula:

$$g(\alpha,\beta) = \sum_y\sum_x f(x,y)h(\alpha-x,\beta-y) \tag{7}$$

Harris Stephens corner detection:

$$\mathcal{H} = \sum_{window}\{(\nabla I)(\nabla I)^T\} \tag{8}$$

$$= \sum_{window}\begin{bmatrix} I_xI_x & I_xI_y \\ I_yI_x & I_yI_y \end{bmatrix}. \tag{9}$$

$$R = \det(\mathcal{H}) - k(\frac{\mathsf{trace}(\mathcal{H})}{2})^2. \tag{10}$$

Rotation matrix, 2D:

$$\mathbf{R}_{2D} = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \tag{11}$$

Camera model:

$$\lambda\mathbf{p} = \mathcal{K}\Pi_0\mathbf{TR}\,{}^W\mathbf{P} = \mathcal{M}\mathbf{P}, \tag{12}$$

Here $\mathbf{p} = [x\ y\ 1]^T$ is the image coordinates in number of pixels and ${}^W\mathbf{P} = [X\ Y\ Z\ 1]^T$ the world coordinates in meter.

$$\mathcal{K} = \begin{bmatrix} \alpha & -\alpha\cot\theta & x_0 \\ 0 & \frac{\beta}{\sin\theta} & y_0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{13}$$

where $\alpha = kf = \frac{f}{\Delta x}$ and $\beta = lf = \frac{f}{\Delta y}$.

$$\Pi_0 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}. \tag{14}$$

$$\mathbf{TR} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \mathcal{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix}, \tag{15}$$

$$\mathcal{M} = \mathcal{K}\Pi_0\mathbf{TR}. \tag{16}$$

$$\mathcal{M} = \mathcal{K}\begin{bmatrix} \mathcal{R} & \mathbf{t} \end{bmatrix}. \tag{17}$$