DATE: 28. Feb. 2017

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 6 PAGES

NOTES: Formulas are found on pages 7 - 8.

Exercise 1

(25%)

a) Give a brief explanation of the following expressions and concepts:

1. point spread function

2. image resolution

3. contrast in an image

b) i) What type of operators or filters can be represented by a filter mask. Explain how the filter mask is used to produce an output image (you can use a sketch).

ii) Show example of filter masks that would blur the image, and sharpen or extract edges of an image.

iii) Does the order of which we apply two linear operators on an image make any difference on the result?

c) A linear and separable operator can be written as:

$$g = h_c^T f h_r \tag{1}$$

Consider a small image:

$$f(x,y) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{pmatrix} \tag{2}$$

i) Design the matrix h_r that will give the average of the pixel position with the next in the row. Let $g_{temp} = fh_r$, What is the content of h_r and g_{temp}?

ii) If we want to do the similar smoothing over the columns, define h_c, and provide the final g.

iii) Comment on the last row and column of the output image g.

d) The definition of 2D Discret Fourier Transform (DFT) and its invers is found in the Formula section. Define matrix U with element: $U(x, \alpha) = \frac{1}{N} e^{\frac{-j2\pi x\alpha}{N}}$, and let $\mathbf{u}_i$ denote a column in U

i) Describe in words (1-2 sentences) what it means that the 2D DFT is separable.

ii) For the image $f(k, l)$, write the DFT, $\hat{f}(m, n)$, as a function of U and $f(k, l)$.

iii) Why is the Fourier transform more commonly used than other transforms?

Exercise 2

(25%)

a) i) Otsus method is much used for image segmentation. Describe the core idea of the method, use figures and equation (10) in the explanation.

ii) Mention other method(s) for segmentation of images.

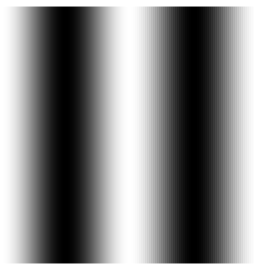
b) Real world images are usually exposed to some kind of noise. Explain some different types of noise, and how we can deal with them.

c) i) Why is Gaussian filters much used in image processing?

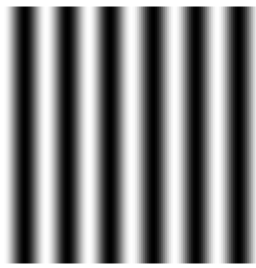
ii)What is the Laplacian of Gaussian (LoG) filter? Please use sketches when explaining.

iii) The LoG filter is often approximated by the difference of Gaussian (DoG) filter. Explain how we can find the DoG filtered output of an image.

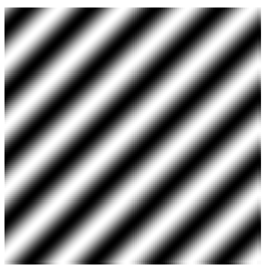
d) Sktech the 2D Fourier transform of the images in Figure 1



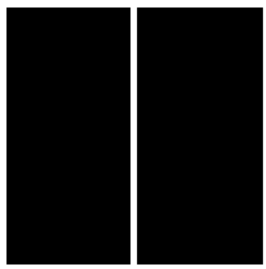
(a)



(b)



(c)



(d)

Figure 1: Figure to problem 2d)

Exercise 3

(25%)

In the front of a locomotive there is three cameras as shown in Figure 2. The angle between the optical axis of the center camera, C, and the world coordinate Z (the horizontal plane) is ϕ degrees. Let $\phi = 7.125$ degrees ($\tan(\phi) = 1/8$). The world coordinates, $\mathbf{P}_w = [X\ Y\ Z]^T$, is placed with origin (center) in the ground plane in front of the locomotive as shown in Figure 2. The camera sensors has 4000×4000 pixels covering an area of $6\,\text{mm} \times 6\,\text{mm}$. There is no skewness. Note that the size of the grid in Figure 2 is in units of 1.0 m. The points $\mathbf{P}^i\ i \in \{1, 2, 3, 4, 5, 6\}$ represents light spots on poles along the railway line.

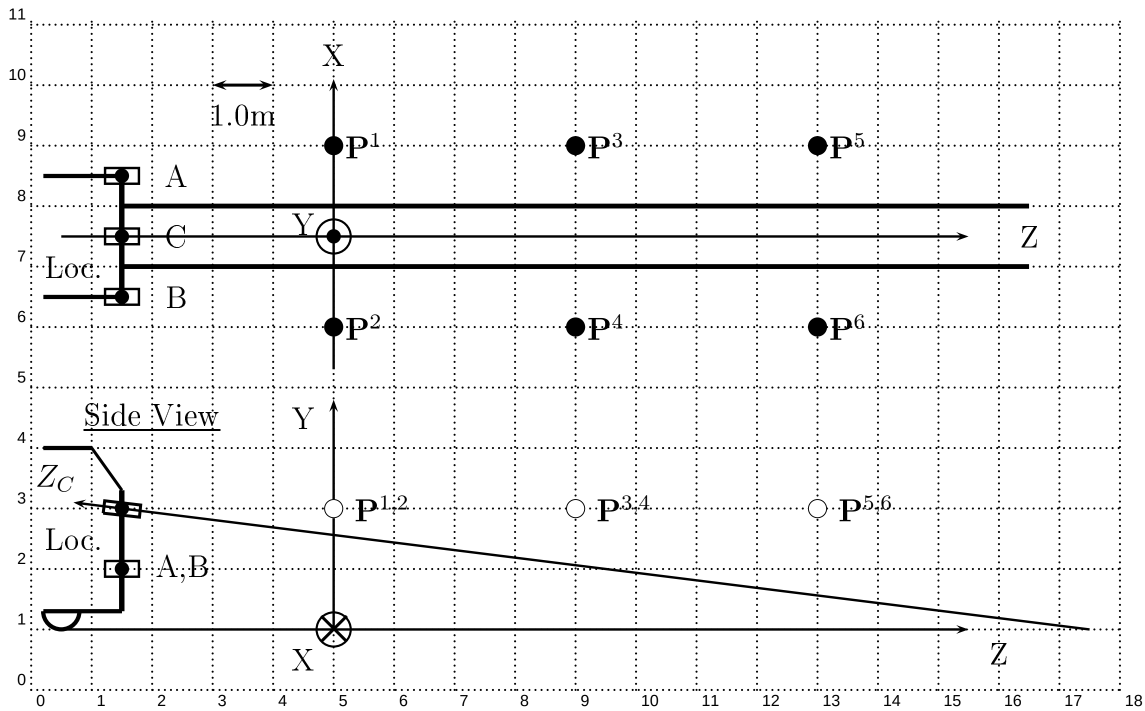


Figure 2: Locomotive, world coordinates and light points.

First we consider the center camera. The camera coordinates are $\mathbf{P}_c = [X_c\ Y_c\ Z_c\ 1]^T$, where X_c is parallel to X. The *field of view* (FOV) given as an angle centered around the optical axis is 40 degrees.

a) Find the **focal length** of the camera.

b) What is the closest point on the ground that can be seen by the camera, given as a world coordinate Z_{min} ($X = Y = 0$)?

c) Find the internal calibration matrix $\mathcal{K}$. (If you have not found a focal length in a) use $f = 9.0\,\text{mm}$.)

The relationship between camera coordinates and world coordinates are given by the following matrix

$$\mathbf{T}R = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & c & -s & t_y \\ 0 & s & c & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{1}$$

where $c = -0.9923$, $s = 0.124$, $t_y = 1.5504\,\mathrm{m}$ and $t_z = -3.721\,\mathrm{m}$.

d) What does the numbers c and s represent, and how are they computed in this case?

e) Compute the camera matrix, $\mathcal{M}$. If you have not found an internal calibration matrix use

$$\mathcal{K} = \begin{bmatrix} 6000 & 0 & 2000 \\ 0 & 6000 & 2000 \\ 0 & 0 & 1 \end{bmatrix}. \tag{2}$$

f) Find the image of the light points $\mathbf{P}^i$ $i \in \{1, 2, 3, 4, 5, 6\}$, the image points $\mathbf{p}^i = [x_i\ y_i]^T$. Are all points within the FOV? What about the y_i for all points?

Exercise 4

(25%)

We will now consider the two cameras, A and B, see Figure 2, used for stereo imaging. The optical axis for camera $\mathbf{A}$ and $\mathbf{B}$ are parallel and the optical axis is horizontal. Then we get the camera matrixes

$$\mathcal{M}^A = \begin{bmatrix} 6000 & 0 & -2000 & -13000 \\ 0 & -6000 & -2000 & -1000 \\ 0 & 0 & -1 & -3.5 \end{bmatrix}, \tag{1}$$

$$\mathcal{M}^B = \begin{bmatrix} 6000 & 0 & -2000 & -1000 \\ 0 & -6000 & -2000 & -1000 \\ 0 & 0 & -1 & -3.5 \end{bmatrix}. \tag{2}$$

In the next questions we look at the images of the points along the line given by $\mathbf{P}_w = [1.5\ 2.0\ Z\ 1]^T$.

a) Show that the image points along the line given above, as a function of the position Z, for camera A can be written as follows:

$$\mathbf{p}_A = \begin{bmatrix} x_A \\ y_A \end{bmatrix} = \begin{bmatrix} 2000 \\ 2000 \end{bmatrix} - \frac{3000}{Z + 3.5} \begin{bmatrix} 1 \\ -2 \end{bmatrix}. \tag{3}$$

Hint: Use $\lambda \mathbf{p} = \mathcal{M}\mathbf{P}_w$ and find λ from the third row.

b) For camera B we have

$$\mathbf{p}_B = \begin{bmatrix} x_B \\ y_B \end{bmatrix} = \begin{bmatrix} 2000 \\ 2000 \end{bmatrix} - \frac{3000}{Z + 3.5} \begin{bmatrix} 5 \\ -2 \end{bmatrix}. \tag{4}$$

Compute the disparity for the two points $\mathbf{P}^i$ $i \in \{1, 3\}$.

c) Define the *epipolare planes* for these points. Where are the *epipolare lines*?

The central camera C is used for estimation of *optical flow*. An object is lying between the rails in front of the train at point $\mathbf{P}_w = [0\ \ 0\ \ Z(t)\ \ 1]^T$, where $Z(t) = 10 - V_z t$. We let the world coordinate system move with the locomotive and the relative speed is $V_z = 20$m/s. The image frame rate is 40 frames per second. The time is then $t = \Delta t \cdot k$ where Δt is the time between each frame. Let the camera matrix for the central camera be

$$\mathcal{M}^C = \begin{bmatrix} 6000 & 250 & -2000 & -7400 \\ 0 & -5700 & -2700 & 1860 \\ 0 & 0.124 & -1 & -3.7 \end{bmatrix}, \tag{5}$$

d) Compute the position of the image point for this object as a function of time. What is the optical flow vector at time $t = 0$?

Formulas

Discrete Fourier transform (DFT) and the invers DFT:

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

$$g(k,l) = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} \hat{g}(m,n)e^{j2\pi[\frac{km}{M}+\frac{ln}{N}]} \tag{7}$$

The 2D convolution formula:

$$g(\alpha,\beta) = \sum_{y}\sum_{x} f(x,y)h(\alpha - x, \beta - y) \tag{8}$$

Let i be illumination function and r reflectance function:

$$f(x,y) = i(x,y)\cdot r(x,y) \tag{9}$$

Between class variance:

$$\sigma_B^2(t) = \frac{[\mu(t) - \bar{\mu}\theta(t)]^2}{\theta(t)(1-\theta(t))} \tag{10}$$

LoG function:

$$LoG = [\frac{x^2+y^2-2\sigma^2}{\sigma^4}]e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{11}$$

$$\mathcal{H} = \sum_{window}\{(\nabla I)(\nabla I)^T\} \tag{12}$$

$$= \sum_{window}\begin{bmatrix} I_x I_x & I_x I_y \\ I_y I_x & I_y I_y \end{bmatrix}. \tag{13}$$

$$\det(\mathcal{H}) - k(\frac{\mathsf{trace}(\mathcal{H})}{2})^2. \tag{14}$$

$$\mathbf{R}_{2D} = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \tag{15}$$

$$\lambda\mathbf{p} = \mathcal{K}\Pi_0\mathbf{T}\mathbf{R}\,{}^{W}\mathbf{P} = \mathcal{M}\mathbf{P}, \tag{16}$$

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.

$$\mathcal{K} = \begin{bmatrix} \alpha & -\alpha \cot\theta & x_0 \\ 0 & \frac{\beta}{\sin\theta} & y_0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{17}$$

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

$$\Pi_0 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}. \tag{18}$$

$$\mathbf{TR} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \mathcal{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix}, \tag{19}$$

$$\mathcal{M} = \mathcal{K}\Pi_0\mathbf{TR}. \tag{20}$$

$$\mathcal{M} = \mathcal{K}\begin{bmatrix} \mathcal{R} & \mathbf{t} \end{bmatrix}. \tag{21}$$