DATE: 26. Feb. 2016

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 7 PAGES

NOTES: Formulas are found on page 8.

Exercise 1

(25%)

a) Explain how a digital grey-level image is represented in a computer. Explain how a digital color image can be represented on a computer.

b) What is the purpose of the stacking operator, given in Equation 1?

$$\mathbf{f} = \sum_{n=1}^{N} N_n f \mathbf{V}_n \tag{1}$$

Let $f = \begin{bmatrix} 4 & 6 & 0 \\ 7 & 5 & 4 \\ 2 & 1 & 1 \end{bmatrix}$ and find $\mathbf{f}$, show all steps.

c) Explain what we mean by a linear operator in an image processing context. Is the stacking operator linear?

d) Equation 2 gives the discrete Fourier transform (DFT) and Eq.3 it's invers.

$$\hat{g}(m,n) = \frac{1}{MN} \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} g(k,l) e^{-j2\pi[\frac{km}{M} + \frac{ln}{N}]} \tag{2}$$

$$g(k,l) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \hat{g}(m,n) e^{j2\pi[\frac{km}{M} + \frac{ln}{N}]} \tag{3}$$

i) Can you explain in a couple of sentences what the Fourier Transform and the discrete Fourier Transfrom of a signal represents? You can use the equations as help if you like.

ii) The Fourier transform is a separable transform. How is this useful?

iii) The convolution theorem is one of the things that makes FT very useful. Explain why?

e) Look at Figure 1. Which frequency specter belongs to which image? Explain briefly why.

Exercise 2

(25%)

a) Explain what we mean by histogram stretching and histogram equalization. Explain the difference. How can we do local histogram equalization? Look at the histogram in figure 2. What would the histogram look like after histogram stretching?

b) When we do filtering in time or frequency domain we need to understand what the filtering means in the other domain.

i) Sketch a box filter and a gaussian filter in the time domain, and roughly sketch or explain in words what you know about their corresponding frequency responses.

ii) Sketch an ideal filter in the frequency domain and roughly sketch or explain in words what you know about the corresponding time response of the filter.

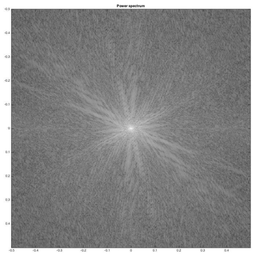
iii) Which type of filter is most preferred, and why?

c) Let the pixels of an object in an image be distributed as: $p_o(x) = \frac{3}{32}(-x^2 + 14x - 45)$ and for the background the pixel values are distributed as: $p_b(x) = \frac{3}{4}(-x^2 + 10x - 24)$. Assume we know that the size of and distance to the object corresponds to the object covering $\frac{2}{3}$ of the image. Find the minimum error threshold.

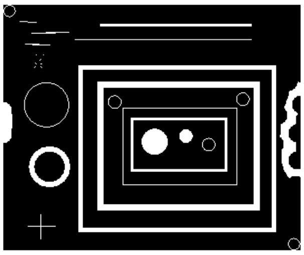
How can you deal with thresholding when the background illumination is nonuniform?

d) Explain the Canny edge detection method.

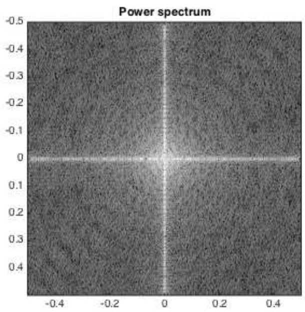
a)



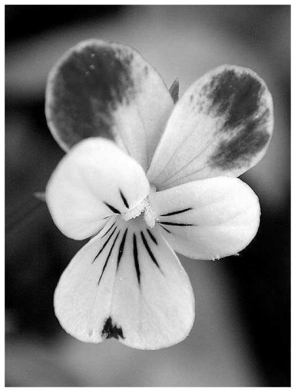
b)



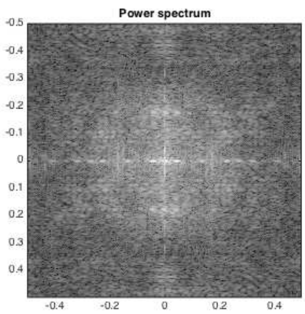
c)


Power spectrum
-0.5
-0.4
-0.3
-0.2
-0.1
0
0.1
0.2
0.3
0.4
-0.4
-0.2
0
0.2
0.4

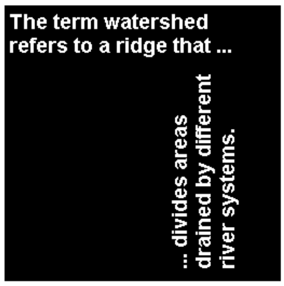
d)



e)


Power spectrum
-0.5
-0.4
-0.3
-0.2
-0.1
0
0.1
0.2
0.3
0.4
-0.4
-0.2
0
0.2
0.4

f)


The term watershed
refers to a ridge that ...
... divides areas
drained by different
river systems.

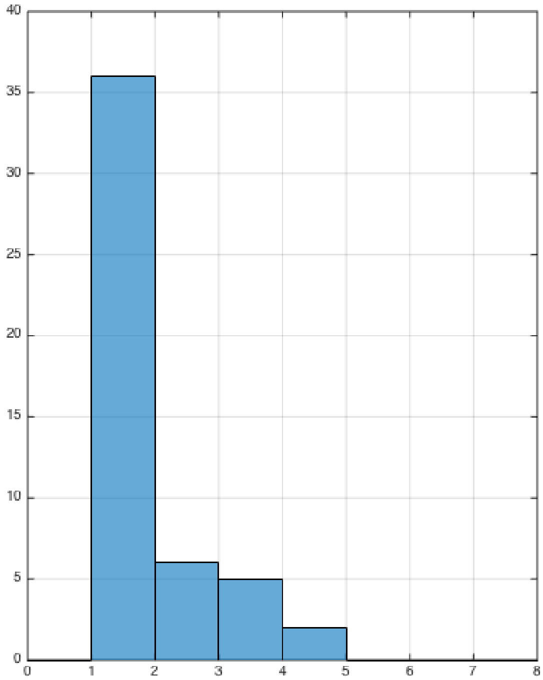


Figure 2: A histogram of the pixel values in a small image.

Exercise 3

(25%)

In Figure 3 a sketch of an autonomous vehicle is shown. The vehicle has 3 cameras, two in the front and one directed at an angle 45 degrees downwards from a rod on the roof of the vehicle. The world coordinates, $\mathbf{P}_w = [X\ Y\ Z]^T$, is placed with origin (center) in the ground plane underneath the vehicle. The camera $\mathbf{C}$ sensor has 4000×4000 pixels covering an area of $6\,\text{mm} \times 6\,\text{mm}$. The two cameras, $\mathbf{A}$ and $\mathbf{B}$, in the front has an imaging sensor with 1600×1600 pixels covering an area of $4\,\text{mm} \times 4\,\text{mm}$. There is no skewness on any of the cameras. Note that the size of the grid in Figure 3 and Figure 4 is in units of 0.5 m.

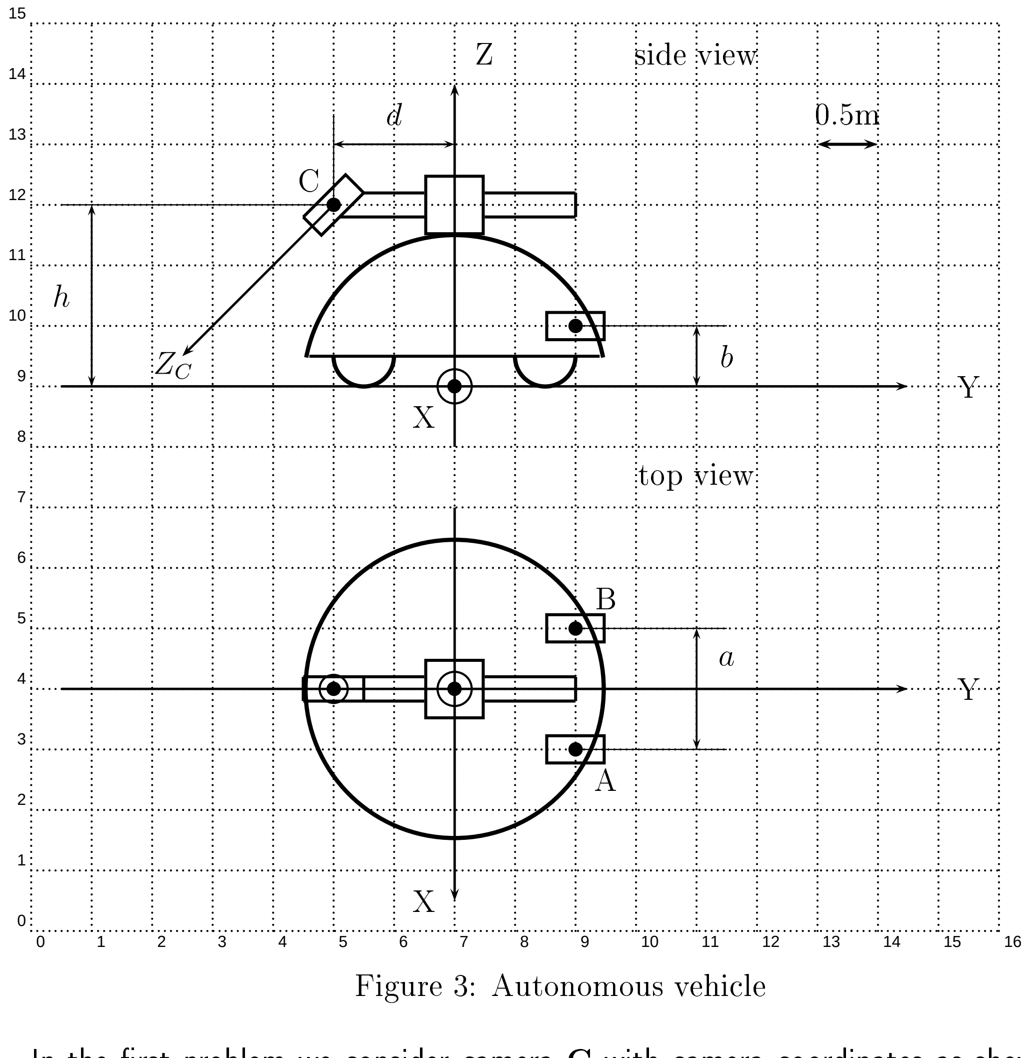


Figure 3: Autonomous vehicle

In the first problem we consider camera $\mathbf{C}$ with camera coordinates as shown in Figure 4. The camera coordinates are $\mathbf{P}_c = [X_c\ Y_c\ Z_c\ 1]^T$, where X_c is parallel to X. The focal length of this camera is $10.0\,\text{mm}$.

a) Find the *field of view* (FOV) given as an angle centered around the optical axis.

b) Show that the relationship between camera coordinates and world coordinates are given by the following matrix

$$\mathbf{R}T = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -a & a & -2.5a \\ 0 & -a & -a & 0.5a \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{1}$$

where $a = \frac{1}{2}\sqrt{2}$ and $\mathbf{P}_c = \mathbf{RTP}_w$. Hint: See Figure 4, rotation β around the X-axis and translation vector $\mathbf{t}$.

c) Find the region in the ground plane covered by the camera FOV. Sketch the region approximately (It is not necessary to compute the exact size of the region. Give a rough sketch with a mark were the optical axis crosses the ground plane.).

d) Find the connection between camera coordinates (not world coordinates) and pixel coordinates for this camera ($\lambda \mathbf{p} = \mathcal{K}\Pi_0 \mathbf{P}_c = \mathcal{M}\mathbf{P}_c$, $\mathbf{p} = [x\ y\ 1]^T$). What is λ in this case?

e) Is it possible in this case to use weak perspective when we want to image small objects on the ground? Explain why or why not.

f) Give a short description of the *Euclidean structure from motion problem.*

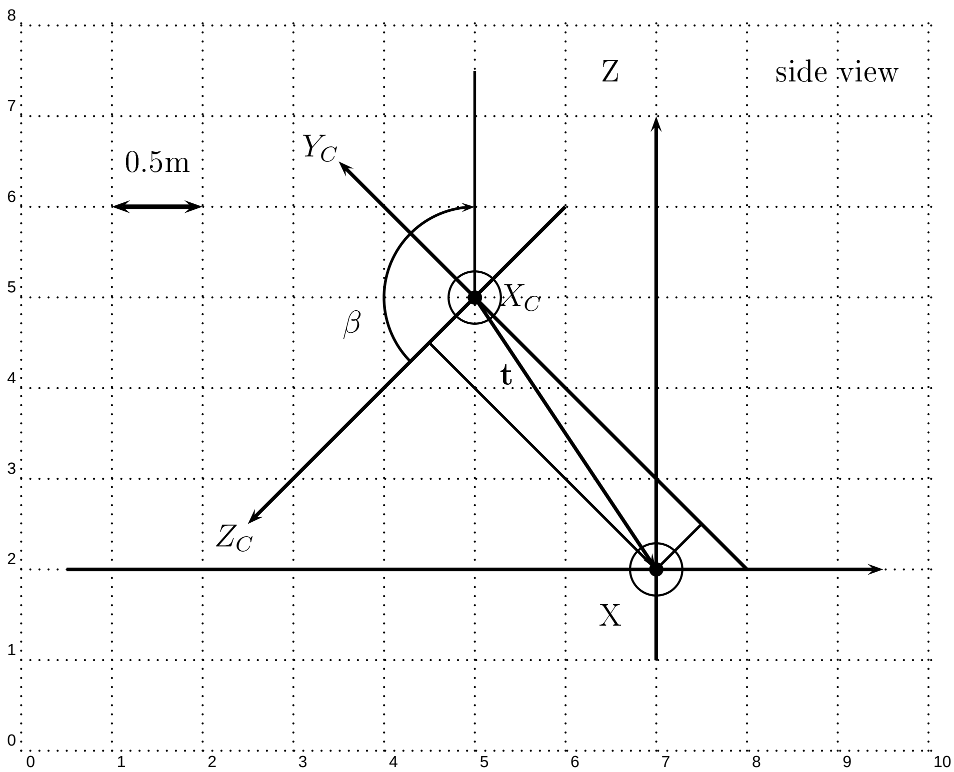


Figure 4: Relationship between Camera coordinates and World coordinates

Exercise 4

(25%)

We will now consider the two cameras in the front of the vehicle. The optical axis for camera $\mathbf{A}$ and $\mathbf{B}$ is parallel. The focal length is $f = 5.0\,$mm. The position of the cameras can be found in Figure 3 with $a = 1.0\,$m and $b = 0.5\,$m. The camera matrixes are

$$\mathcal{M}^A = \begin{bmatrix} -2000 & -800 & 0 & 1800 \\ 0 & -800 & 2000 & -200 \\ 0 & -1 & 0 & 1 \end{bmatrix}, \tag{1}$$

$$\mathcal{M}^B = \begin{bmatrix} -2000 & -800 & 0 & -200 \\ 0 & -800 & 2000 & -200 \\ 0 & -1 & 0 & 1 \end{bmatrix}. \tag{2}$$

In the next questions we look at the images of the points along the line given by $\mathbf{P}_w = [0\ Y\ 0\ 1]^T$.

a) Find the image points in the two cameras for the world points given above as a function of the position Y. What is the closest point in front of the vehicle that can be seen by both cameras?

b) Compute the disparity for the points above. What is the distance, Y, given by the disparity (pixel coordinates in the two cameras)?

c) Define the *epipolare plane* for this case.

The two cameras are rotated around their camera Z-axis such that their optical axis meet in the point $\mathbf{P}_w = [0\ 3.0\,\text{m}\ 0.5\,\text{m}\ 1]^T$.

d) What are the position of the *epipoles* for this stereo system? Hint: Make a sketch of the plane formed by the base line and the optical axis of the cameras and use simple geometric relationships in the computation. Use normalized image plane units.

Formulas

$$\mathcal{H} = \sum_{window} \{(\nabla I)(\nabla I)^T\} \tag{3}$$

$$= \sum_{window} \begin{bmatrix} I_x I_x & I_x I_y \\ I_y I_x & I_y I_y \end{bmatrix}. \tag{4}$$

$$\det(\mathcal{H}) - k(\frac{\mathsf{trace}(\mathcal{H})}{2})^2. \tag{5}$$

$$\mathbf{R}_{2D} = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \tag{6}$$

$$\lambda \mathbf{p} = \mathcal{K}\Pi_0 \mathbf{TR}\,{}^W\mathbf{P} = \mathcal{M}\mathbf{P}, \tag{7}$$

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.

$$\mathcal{K} = \begin{bmatrix} \alpha & -\alpha\cot\theta & x_0 \\ 0 & \frac{\beta}{\sin\theta} & y_0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{8}$$

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

$$\Pi_0 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}. \tag{9}$$

$$\mathbf{TR} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \mathcal{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix}, \tag{10}$$

$$\mathcal{M} = \mathcal{K}\Pi_0\mathbf{TR}. \tag{11}$$

$$\mathcal{M} = \mathcal{K}\begin{bmatrix} \mathcal{R} & \mathbf{t} \end{bmatrix}. \tag{12}$$