DATE: 4. Dec. 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 5 PAGES

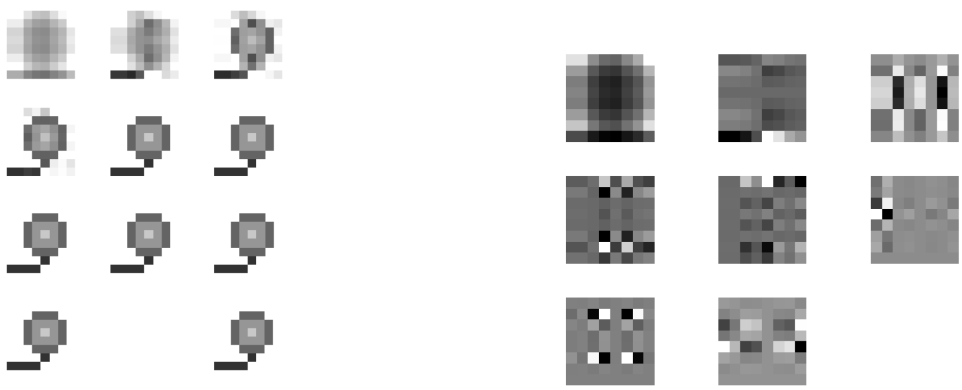
NOTES: Formulas are found on pages 6 - 7.

Exercise 1

(25%)

a) Explain briefly the following:

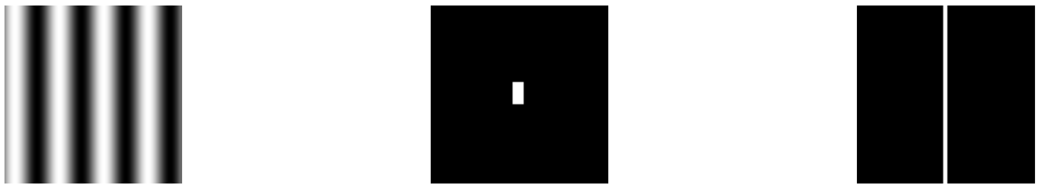
1) Image resolution and image quantization.
2) What do we mean by a separable operator?
3) What is the stacking operator?



(a) (b)

Figure 1: Figure to problem 1b)

b) Figure1 a) shows the approximation of the flower in the bottom right corner by SVD decomposition. In figure b) we see the basis images.
Explain how to interpret these figures related to the equation $f = U\Lambda^{\frac{1}{2}}V^T$



(a) (b) (c)

Figure 2: Figure to problem 1c)

c) Figure 2 shows three different images. Sketch the corresponding magnitude respons of the Discrete Fourier Transform (DFT) of the images, and include a brief explanation.

d) Explain why the discrete histogram equalization in general does not give a flat histogram. Use Equation 7 in your explanation.

Exercise 2

(25%)

a) Different types of noise require different filtering strategies. Explain what we mean by Gaussian noise and salt and pepper noise, and explain a filter mask strategy that would be good in each of the noise cases.

b) Let I be an image.

$$I = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}, \tag{1}$$

Find an output image that is the derivative over the rows and smooth over columns by using a Prewitt filtermask. Explain in words how you find the Prewitt gradient image.

c) The Laplacian can be computed using the following mask

$$\mathbf{h}_L = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}. \tag{2}$$

Consider a 1D discrete signal, $\cdots f(n-1)\ f(n)\ f(n+1) \cdots$. Use simple differences to compute the differential, $f'(n) = f(n+1) - f(n)$

i) Find the second derivative for the 1D signal above and show how this leads to the filter mask in (2) when the effect from the two directions are added together.
ii) LoG - Laplacian of Gaussian filter is often used. Explain briefly what a LoG filter is and why it is beneficial compared to the Laplacian.

d) You want to distinguish a brighter foreground object from a darker background. You have reasons to believe that the object occupies approximately $\frac{8}{9}$ of the image. Let the probability density function of the intensity values of the object and background be defined as:

$$p_o(t) = \frac{3}{32}(-t^2 + 14t - 45) \tag{3}$$

$$p_b(t) = \frac{3}{4}(-t^2 + 10t - 24) \tag{4}$$

respectively. Find the optimal threshold by minimizing the number of misclassified pixels.

Exercise 3

(25%)

An airplane is flying straight ahead in height h above the ground. A digital camera is placed under the plane with optical axis vertically, imaging objects on the ground. Assume that the ground is completely flat (ignore the curvature of the earth). Let the world coordinates be defined with the Z axis vertical upwards with origin in the ground plane. At time $t = 0$ the optical center of the camera is exactly vertically above the origin of the world coordinates. Let the world X coordinate be parallel to the camera coordinate X_c. The homogeneous camera coordinates are $\mathbf{P}_c = [X_c\ Y_c\ Z_c\ 1]^T$. The focal length of the camera is $f = 40\,\text{mm}$ and the image sensor has 2000×4000 pixels in an area $10\,\text{mm} \times 20\,\text{mm}$. Zero skew. The airplane is heading in the X-direction (constant translation).

a) Make a 2D sketch of the situation above, use the X-Z-plane and let $Y = 0$.

b) Find the field of view (FOV), given by angles in the $Z_c - X_c$ and $Z_c - Y_c$-plane. How large is the area covered on the ground plane?

c) Find the connection between camera coordinates and world coordinates given by the rotation matrix $\mathbf{R}$ and the translation vector $\mathbf{t}$.

d) Find the internal calibration matrix $\mathcal{K}$.

e) Show that the normalized ($m_{34} = 1$) camera matrix is given by the following expression, where the height h is a parameter.

$$\mathcal{M} = \begin{bmatrix} -\frac{8000}{h} & 0 & -\frac{2000}{h} & 2000 \\ 0 & -\frac{8000}{h} & -\frac{1000}{h} & 1000 \\ 0 & 0 & -\frac{1}{h} & 1 \end{bmatrix}. \tag{1}$$

A building with rectangular base is placed with one corner in the origin. The length of the building is $L = 100\,\text{m}$, the width is $W = 50\,\text{m}$ and the height is $H = 25\,\text{m}$. Let the flying height be $h = 1000\,\text{m}$.

f) Make a sketch of the building in the image plane. Because $h >> H$ we can use $Z \approx 0$ for all corners of the building.

Exercise 4

(25%)

Assume that you have 2 cameras of the same type as in exercise 3. You will use these for stereo imaging.

a) Make a sketch where the cameras are placed with distance t between the camera centers and the optical axis forms an angle $\alpha < 90$ degrees. Define the *epipolare plane* and the *epipoles*.

b) Let the optical axis above be parallel ($\alpha = 0$ degrees). Choose a scene point that is imaged in the two image planes and show with reference to the image points how the *disparity* is defined (no computation has to be done).

c) Simplify the case in b) by choosing a 2D view (let the perpendicular camera coordinate be zero). Compute the disparity for one direction, parallel to the base line. Show how the *disparity* can be used to compute the depth (distance from the camera centers), Z.

We will now go back to the airplane in exercise 3. The plane is moving in the X-direction with constant height h. The velocity of the plane is V_x. Let the flying height be $h = 1000\,$m. The camera matrix can then be written:

$$\mathcal{M} = \begin{bmatrix} -8 & 0 & -2 & 2000 - 8V_x t \\ 0 & -8 & -1 & 1000 \\ 0 & 0 & -0.001 & 1 \end{bmatrix}. \tag{1}$$

d) How will the image points for scene points on the ground move as a function of time? What is the optical flow? Use $\mathbf{P}_w = [X \quad Y \quad 0 \quad 1]$ and find the corresponding point in the image plane as a function of time. (We ignore the real height of objects on the ground, $h >> H$.)

The camera captures images with a frame rate of 100 frames per second.

e) Find the velocity vector, *optical flow vector*, at $t = 0.1\,$s, when $V_x = 100$m/s, $V_y = V_z = 0$.

Formulas

SVD decomposition

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

Discrete Fourier transform (DFT) and the invers DFT:

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

$$g(k,l) = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} \hat{g}(m,n)e^{j2\pi[\frac{km}{M}+\frac{ln}{N}]} \tag{4}$$

The 2D convolution formula:

$$g(\alpha,\beta) = \sum_{y}\sum_{x} f(x,y)h(\alpha - x, \beta - y) \tag{5}$$

Let i be illumination function and r reflectance function:

$$f(x,y) = i(x,y)\cdot r(x,y) \tag{6}$$

histogram equalization

$$s = \lceil G \cdot \sum_{g=g_{min}}^{R} p_{old}(g) - 1 \rceil \tag{7}$$

h and h^T represent Prewitt and Sobel filters by using different K's.

$$h = \begin{bmatrix} -1 & 0 & 1 \\ -K & 0 & K \\ -1 & 0 & 1 \end{bmatrix}, \tag{8}$$

Between class variance:

$$\sigma_B^2(t) = \frac{[\mu(t) - \bar{\mu}\theta(t)]^2}{\theta(t)(1-\theta(t))} \tag{9}$$

LoG function:

$$LoG = [\frac{x^2+y^2-2\sigma^2}{\sigma^4}]e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{10}$$

Minimum errror threshold:

$$\theta p_o(t) = (1-\theta)p_b(t) \tag{11}$$

Harris Stephens corner detection:

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

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

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

Rotation matrix, 2D:

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

Camera model:

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

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 $\beta = lf = \frac{f}{\Delta y}$.

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

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

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

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