DATE: 7. Dec. 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 9 PAGES

NOTES: Formulas are found on page 10-11.

Exercise 1

(25%)

a) Give a brief explanation of what we mean by the following expressions and some examples of why we do it:

1. image coding and compression

2. image enhancement

3. image segmentation

4. image description

b) The color image *peppers* from Matlab is shown in Figure 1. In Figure 2 you see a figure with different color space mappings of the image.
i) Explain briefly how color images are represented in a computer.
ii) Identify the R, G, and B images as well as the luminance image from Figure 2, and explain how you identified them.

c) The discrete Fourier transform (DFT) and the invers DFT is written as:

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

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

Using Eq (1), find an expression for $\hat{g}(0,0)$ of an image $g(x,y)$ and explain what it means.

Figure 1: Figure to problem 1 b) - peppers color image.

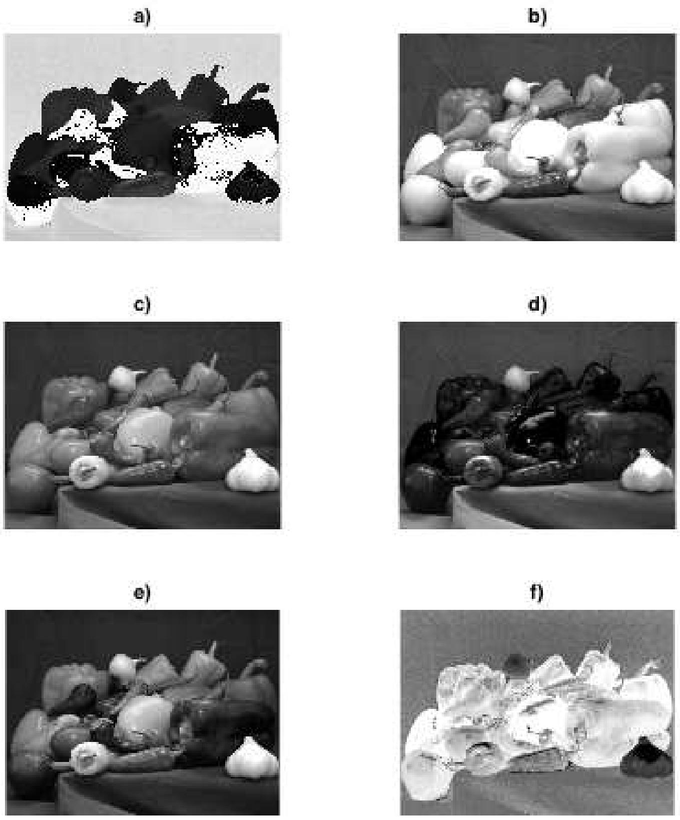

Figure 2: Figure to problem 1 b) - peppers color space images.

d) Let the input and output images to an operator be called f and g respectively.

Imagine the operator can be described by a separable and shift invariant point spread function, which can be written as:

$$h(x, \alpha, y, \beta) = h_c(x - \alpha)h_r(y - \beta) \tag{3}$$

Explain how the point spread function of a linear operator can be interpreted by defining x, y relative to α, β using f and g. Rewrite the equation if the point spread function was separable but not shift invariant. Rewrite the equation if the point spread function was shift invariant but not separable.

Exercise 2

(25%)

a) Define

$$h_r(\beta - y) = \begin{cases} 1 & \text{if } \beta - y = -1 \\ 2 & \text{if } \beta - y = 0 \\ 3 & \text{if } \beta - y = 1 \end{cases} \tag{1}$$

$$h_c(\alpha - x) = \begin{cases} 1 & \text{if } \alpha - x = 0 \\ 1 & \text{if } \alpha - x = 1 \end{cases} \tag{2}$$

and look at a (very small) image matrix:

$$f = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 2 & 4 \\ 1 & 3 & 5 \end{bmatrix}. \tag{3}$$

Find the 3×3 output image g,and show how you found it by showing some of the calculation steps. You can choose (but do not have to) use a matrix formulation, but then you have to show all involved matrices.

b) Look at Figure 3 and the following known filtermasks:

$$h_1 = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix}. \tag{4}$$

$$h_2 = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}. \tag{5}$$

i) Let figure 3 a) be the original image. which of the images corresponds to filtering with filtermask h_1 and h_2? Explain why.

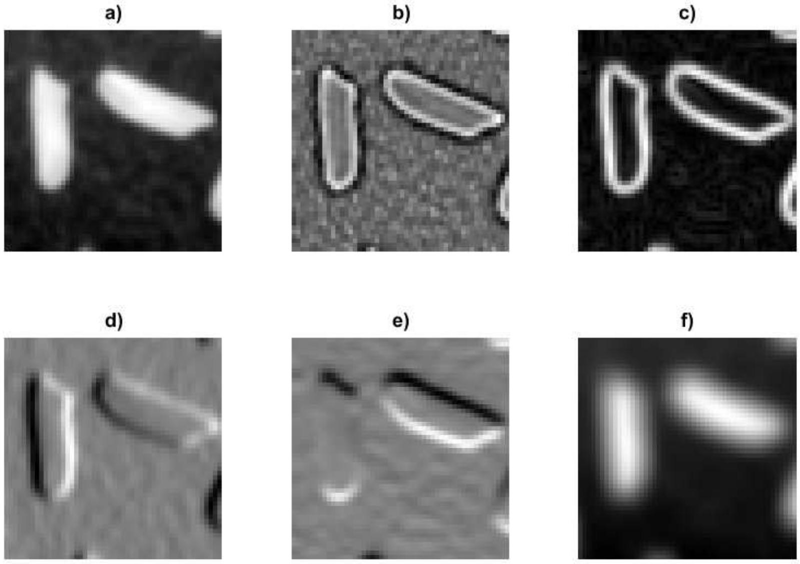

Figure 3: Figure to problem 2 b).

ii) Consider an image with smoothed, slow rising, mostly horizontal, edges. Which of these two filtemasks would you have preferred (and why) if you were to find the edges?

c) What is homomorphic filtering? Use figures and/or equations in your explanation. Explain how it can be useful when we have uneven illumination due to the distance from the light source.

d) In figure 4 you see the *cameraman* image with corner points detected by Harris-Stephens corner detector. Explain how the following equation is important for the main idea of the Harris-Stephens detector, and define the symbols in the equation:

$$E(u,v) = \sum_{x,y} w(x,y)[I(x+u,y+v) - I(x,y)]^2 \tag{6}$$

$$E(u,v) \approx [u \;\; v]\, \mathcal{H}\, [u \;\; v]^T \tag{7}$$

Other important Harris-Stephens detector equations are seen in Equations (11) - (13). Can you briefly outline the Harris-Stephens corner detector algorithm (i.e. how are the green corners in the image produced)?

Figure 4: Figure to problem 2 d) - cameraman with corner points.

Exercise 3

(25%)

A sketch of an Unmanned Ariel Vehicle (UAV) is shown in Figure 5. The vehicle has 3 cameras, directed downwards. Camera *a* and *c* is rotated an angle of $\pm\phi$ degrees as shown in the Figure. The camera in the middle, *b*, is directed vertical downwards. The world coordinates, $\mathbf{P}_w = [X\ Y\ Z]^T$, is placed with origin (center) in the ground plane underneath the vehicle. The cameras have image sensors with 1000×1000 pixels covering an area of 6 mm $\times$ 6 mm. The camera centers are positioned on a horizontal line, at $\mathbf{P}_0^a = [0\,\text{m}\ -0.05\,\text{m}\ 1.0\,\text{m}]^T$, $\mathbf{P}_0^b = [0\,\text{m}\ 0\,\text{m}\ 1.0\,\text{m}]^T$ and $\mathbf{P}_0^c = [0\,\text{m}\ 0.05\,\text{m}\ 1.0\,\text{m}]^T$ in world coordinates as given by Figure 5. The camera coordinates are defined as shown in Figure 6. We assume no skewness on any of the cameras. Note that the size of the grid in Figure 5 is in units of 0.1 m.

The camera angle is $\phi = 26.565\,^\circ$ ($\tan(\phi) = 0.5$).

a) What is the **focal length** when the *field of view* (FOV) is 45 degrees?

b) The FOV of camera *a* and *b* overlap in the ground plane. Consider the line given by $\mathbf{P}_w = [0\ Y\ 0]^T$. Find the interval for coordinate Y where the two FOV's overlap.

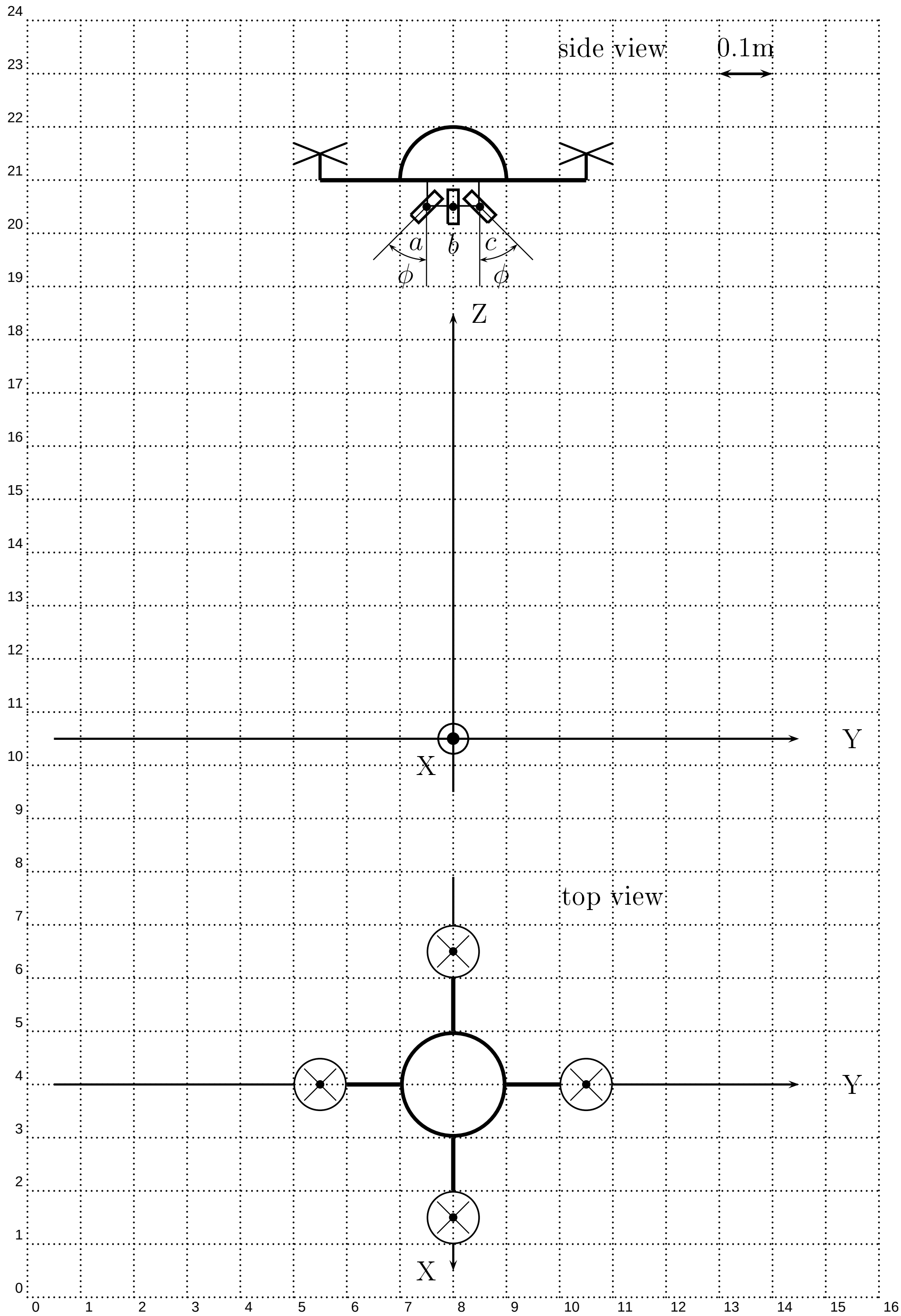

Figure 5: Unmanned Ariel Vehicle

Camera a and b is used for stereo vision. The camera coordinates of b is used as reference (world coordinates). The point imaged by the two cameras is given by $\mathbf{P}_b$ in camera b and $\mathbf{P}_a$ in camera a. The relationship between these two are

$$\mathbf{P}_b = \mathcal{R}\mathbf{P}_a + \mathbf{t}, \tag{1}$$

or the inverse

$$\mathbf{P}_a = \mathcal{R}^T(\mathbf{P}_b - \mathbf{t}). \tag{2}$$

c) Find the rotation matrix $\mathcal{R}$ and the translation vector $\mathbf{t}$. (Note that the translation here is given in camera b coordinates.). Check the result by setting $\mathbf{P}_b = [0 \;\; -0.25\,\mathrm{m} \;\; -1.0\,\mathrm{m}]^T$ and find $\mathbf{P}_a$. Is the result reasonable (use Figure 5 and sketch the result).

d) What is the position of the *epipoles* for this stereo configuration? Make a sketch (see Figure 6) and find the answer in metric numbers.

e) Show that the internal calibration matrix is given by

$$\mathcal{K} = \begin{bmatrix} 1207.1 & 0 & 500 \\ 0 & 1207.1 & 500 \\ 0 & 0 & 1 \end{bmatrix}. \tag{3}$$

f) Find the camera matrix for camera b.

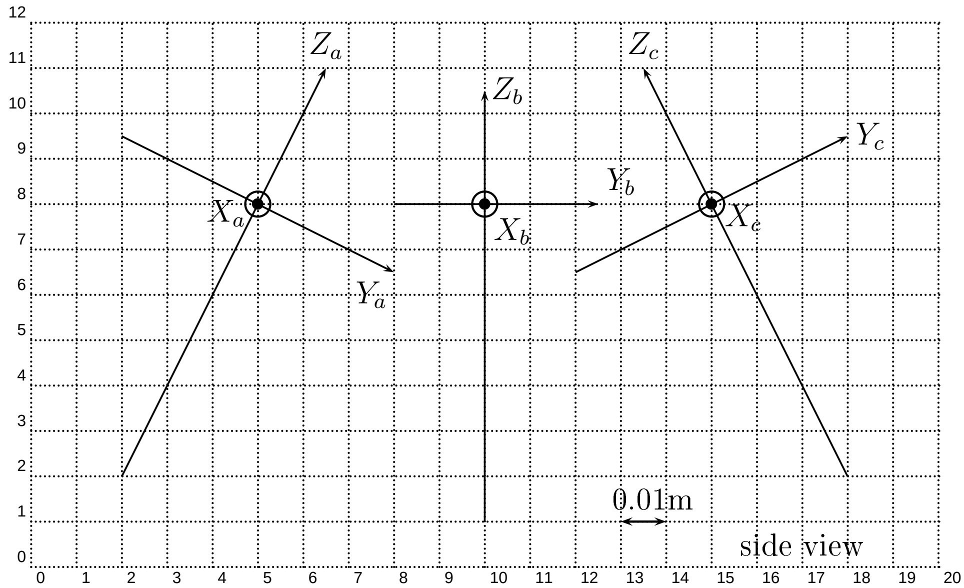

Figure 6: Definition of Camera coordinates

Exercise 4

(25%)

We will now consider camera *b* for the UAV in motion. The focal length is now adjusted to $f = 12.0\,\text{mm}$ and the hight from ground is $2.0\,\text{m}$. The camera matrix is then:

$$\mathcal{M}^b = \begin{bmatrix} -1000 & 0 & -250 & 500 \\ 0 & -1000 & -250 & 500 \\ 0 & 0 & -0.5 & 1 \end{bmatrix}. \qquad (1)$$

The ground underneath the UAV is covered with tiles of size $0.2\,\text{m} \times 0.2\,\text{m}$. The crossing points between the tiles are given by $X = 0.2 \cdot i$, $Y = 0.2 \cdot j$ and $Z = 0$, where $i, j \in \{\cdots -3, -2, -1, 0, 1, 2, 3 \cdots\}$. We now want to find the *optical flow* for these crossing points in the image plane.

First the UAV is moving horizontally with velocity $V_y = 0.5$m/s, ($V_x = V_z = 0$). We consider the row of crossing points along the Y -axis, i.e. $X = 0$. Each point move to the left with position as a function of time given by

$$Y_{crosspoint}(t) = -V_y t + 0.2 \cdot j, \qquad j \in \{\cdots -2, -1, 0, 1, 2, \cdots\}. \qquad (2)$$

The image frame rate is 100 frames per second ($t = 0.01 \cdot k, \quad k \in \{0, 1, 2, \cdots\}$).

a) Compute the position of the crossing points in the image plane and find the **optical flow** for these points.

b) What is the velocity vector for other points in the ground plane? Explain!

Next we consider the UAV starting in the original position ($t = 0$ above) and moving with velocity $V_z = 10$m/s, ($V_x = V_y = 0$) straight upwards. Consider the same crossing points as in the first part, $\mathbf{P}_w = [0.2 \cdot i, \;\; 0.2 \cdot j, \;\; -V_z t, \;\; 1]^T$.

c) Show that the image coordinates for the crossing points as a function of time can be written as (Hint: $\lambda = 0.5V_z t + 1$, third line!)

$$x = 500 - \frac{200 \cdot i}{0.5V_z t + 1} \qquad (3)$$

$$y = 500 - \frac{200 \cdot j}{0.5V_z t + 1} \qquad (4)$$

Use the two first frames to compute the velocity vectors at time $t = 0$.

d) Find the optical flow for the points $(i, j) \in \{(0,0),(1,0),(1,1)\}$ and explain how the optical flow field will look like in this case.

e) Explain shortly what the characteristics of a Robot Vision system is, compared to a Machine Vision system.

Formulas

Discrete Fourier transform (DFT) and the invers DFT:

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

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

The 2D convolution formula:

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

Let i be illumination function and r reflectance function:

$$f(x,y) = i(x,y)\cdot r(x,y) \tag{8}$$

Between class variance:

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

LoG function:

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

$$\mathcal{H} = \sum_{window}\{(\nabla I)(\nabla I)^T\} \tag{11}$$

$$= \sum_{window}\begin{bmatrix} I_xI_x & I_xI_y \\ I_yI_x & I_yI_y \end{bmatrix}. \tag{12}$$

$$R = \det(\mathcal{H}) - k(\frac{\text{trace}(\mathcal{H})}{2})^2. \tag{13}$$

$$\mathbf{R}_{2D} = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \tag{14}$$

$$\lambda \mathbf{p} = \mathcal{K}\Pi_0 \mathbf{TR}\,{}^W\mathbf{P} = \mathcal{M}\mathbf{P}, \tag{15}$$

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.

$$\mathcal{K} = \begin{bmatrix} \alpha & -\alpha\cot\theta & x_0 \\ 0 & \frac{\beta}{\sin\theta} & y_0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{16}$$

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

$$\Pi_0 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}. \tag{17}$$

$$\mathbf{TR} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \mathcal{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix}, \tag{18}$$

$$\mathcal{M} = \mathcal{K}\Pi_0\mathbf{TR}. \tag{19}$$

$$\mathcal{M} = \mathcal{K}\begin{bmatrix} \mathcal{R} & \mathbf{t} \end{bmatrix}. \tag{20}$$

$$\mathbf{v} = \begin{bmatrix} v_x \\ v_y \end{bmatrix} = \frac{T_z}{Z}\begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix} = \frac{T_z}{Z}(\mathbf{p} - \mathbf{p}_0), \tag{21}$$

where

$$\mathbf{p}_0 = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} = \frac{f}{T_z}\begin{bmatrix} T_x \\ T_y \end{bmatrix}. \tag{22}$$

$$\mathbf{v} = -\frac{f}{Z}\begin{bmatrix} T_x \\ T_y \end{bmatrix}. \tag{23}$$