DATE: 9. December 2019

University of Stavanger

EXAM IN: ELE510 IMAGE PROCESSING with COMPUTER VISION

DURATION: 4 hours

ALLOWED REMEDIES: Defined, simple calculator permitted.

THE SET OF EXERCISES CONSISTS OF 5 EXERCISES ON 4 PAGES

NOTES: All subproblems are weighted equally.
Formulas are found on page 5-6.

Exercise 1

a) Explain how a digital image, $I(x, y)$, is represented in a computer. The answer should include the following: sampling, quantization, color image. Feel free to use drawings and/or mathematical expressions in your answer.

b) Sketch a pinhole camera. Mark the focal length, the image plane, virtual image plane, and optical axis and focal point. Use your sketch and and explain how perspective projection forms an image.

Exercise 2

a) Let $I(x, y)$ denote a gray level image of size $N \times M$, and $I'(x, y)$ the corresponding image after a transformation. What is a geometric transformation? Set up the equation for a flop (mirror) operation. Can you briefly explain the difference between forward mapping and inverse mapping in this context?

b) Let I_b be an image.

$$I_b = \begin{bmatrix} 0 & 1 & 2 & 5 \\ 1 & 0 & 1 & 7 \\ 1 & 2 & 6 & 7 \\ 0 & 1 & 6 & 6 \end{bmatrix}, \tag{1}$$

find the *normalized histogram* and the *cumulative distribution function* (CDF) of the image.

c) Explain the purpose of *histogram equalization*. Feel free to use figures as you explain.

Perform histogram equalization on I_b by help of the CDF from b), and show the output image.

d) *Frame differencing* is a simple method for detecting moving object in video frames. The simplest approach is to compute a difference image:

$$I'(x, y) = |I_t(x, y) - I_{t-1}(x, y)| > \tau \tag{2}$$

Make a simple sketch of 3 consecutive frames in a video with a moving ball on an even background. Illustrate the output of Eq. 2. What is the well-known problem of this approach? Propose another well-known approach to solve that problem, and sketch the output.

Exercise 3

a) Let $T(\cdot)$ be a system (filter), and $I(x, y)$ an image; The output image can be written as $I'(x, y) = T(I(x, y))$. Define what we mean by: *linear* system, *shift invariant* system, FIR and IIR?

b) 2D Convolution is shown below with notation from the book:

$$I'(x, y) = I(x, y) * G(x, y) = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} I(x + \tilde{w} - i, y + \tilde{h} - j)G(i, j) \tag{1}$$

where $\tilde{w}, \tilde{h}$ indicates the half width and height of a filter kernel, G.

Using the expressions from a), *when* can we use *convolution* to find $I'(x, y)$? Write up the kernel, G(i,j), of a 3×3 mean filter, and explain with a few sentences what it does.

c) Some 2D filter-kernels are separable. What is the advantage of separable filter kernels? Let the following describe a 2D convolution of an image $I(x, y)$ with a separable Gaussian filter:

$$I'(x, y) = I(x, y) * \left(\frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} \right) = .. \tag{2}$$

Write the decomposed version of the equation.

d) Explain the purpose of the three filter-kernels below. Perform the convolution $I'_b(x, y) = I_b(x, y) * h_2$ (I_b from Eq. 1). Let $I'_b(x, y)$ have the same size as the input $I_b(x, y)$, and use zero padding.

$$\mathbf{h}_1 = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}. \tag{3}$$

$$\mathbf{h}_2 = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}. \tag{4}$$

$$\mathbf{h}_3 = \frac{1}{64}\begin{bmatrix} 1 & 6 & 1 \\ 6 & 36 & 6 \\ 1 & 6 & 1 \end{bmatrix}. \tag{5}$$

e) Equation Eq. 6 shows the magnitude of an ideal 1D low pass filter in the frequency domain, with cutoff freq. of w_0.

$$|H(w)| = \begin{cases} 1 & \text{if} |w| < w_0 \\ 0, & \text{otherwise} \end{cases} \tag{6}$$

Sketch roughly how such an ideal filter in frequency domain appear in the *spatial domain*. Use the formula for the invers discrete time Fourier Transform of a 1D signal to find a mathematic expression for the same:

$$h(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} H(w)e^{jwn}dw \tag{7}$$

Hint: Use Eulers formula (formula pages, F. 2)

Exercise 4

a) A small (7x7) image is depicted in Figure 1 with pixel intensity values. The circle is roughly pointing out a foreground object, this is not a part of the image itself. We wish to do segmentation to get the foreground object. Explain thresholding by using this image as example; Let us use p-tile method for finding the threshold, and from the size of the circle we can say that the object should be approximately 12 pixels. Sketch the output image. Can you mention other ways of finding the threshold?

0	1	0	1	1	2	0
1	1	0	3	4	1	0
0	2	4	5	4	0	1
0	0	4	5	4	1	1
1	0	3	3	2	0	2
0	1	2	0	1	0	5
1	0	1	0	0	3	4

Figure 1: Figure to problem 4 a) and b)

b) Let us look at the image from Figure 1 again and try segmentation by region growing instead. Let the seed point be column=4, row=3, (intensity level =

5). Let our feature, $f(x, y)$, be the pixel intensity value, and the similarity measure be defined $d = |f(x, y) - \bar{f}|$, where $\bar{f}$ denotes the mean of the feature over the region. Include a point in the region if $d < 2$, and use a 4NB (neighbourhood). Go through entire boundary before $\bar{f}$ is updated (as done in class). Do the segmentation, Show the development of $\bar{f}$ in each round, and sketch the final output image. Compare with a) and comment.

c) The Harris Stephens corner detector is a classical algorithm for finding corners, or feature points, in an image. Can you explain minimum two applications where finding such feature points are is necessary or useful? Mention some *properties* that are good for a feature point detector.

Exercise 5

A camera has a focal length of $f = 30mm$ and the image sensor have 3000 x 6000 pixels covering an area of 10x20 mm. The camera has zero skew. The camera is facing a wall, perpendicular to the optical axis, at a distance of 2 meters measured along the optical axis from the focal point. Let $\{X_c, Y_c, Z_c\}$ correspond to the camera coordinate system.

a) What is the Field of View (FOV) in the X_c and Y_c direction (i.e. given by angles in the $Z_c - X_c$ and $Z_c - Y_c$ planes)? How large area on the wall (height and width in m) is covered by the image?

b) Find the internal camera calibration matrix.

c) Assume you have two cameras of the same type, and you will use them for stereo imaging. Let the cameras be *rectified*, i.e. the cameras are translated to each other parallell to the scan lines (rows). Explain what we mean by corresponding points in the images, and how these can be used to find the depth, i.e. distance, to objects in the image scene. What are epipolar lines in the context of stereo imaging. (brief explanations)

d) Let the cameras be rectified, and let the maximum disparity be 20. Which of the following pixel coordinates (x,y) in the right image could not match the pixel (52,3) in the left image, and why : i) (26,3) ii) (48,13), iii) (64,3) iv) (48,3) v) (59,6)

Formulas

Gaussian function

$$Gauss_{\sigma^2}(x,y) = \frac{1}{2\pi\sigma^2} \cdot e^{\left(-\frac{x^2+y^2}{2\sigma^2}\right)} \tag{1}$$

Eulers formula

$$e^{jwn} = cos(wn) + jsin(wn) \tag{2}$$

SVD decomposition

$$f = U\Lambda^{\frac{1}{2}}V^T \tag{3}$$

Discrete Fourier transform (DFT) and the invers DFT:

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

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

The 2D convolution formula:

$$g(\alpha,\beta) = \sum_y \sum_x f(x,y)h(\alpha - x, \beta - y) \tag{6}$$

Let i be illumination function and r reflectance function:

$$f(x,y) = i(x,y) \cdot r(x,y) \tag{7}$$

Between class variance:

$$\sigma_B^2(t) = \frac{[\mu(t) - \bar{\mu}\theta(t)]^2}{\theta(t)(1-\theta(t))} \tag{8}$$

LoG function:

$$LoG = [\frac{x^2+y^2-2\sigma^2}{\sigma^4}]e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{9}$$

Minimum errror threshold:

$$\theta p_o(t) = (1-\theta)p_b(t) \tag{10}$$

Harris Stephens corner detection:

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

$$= \sum_{window} \begin{bmatrix} I_x I_x & I_x I_y \\ I_y I_x & I_y I_y \end{bmatrix}. \tag{12}$$

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

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

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

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

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

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

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

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