School of Computer Science and Communication, KTH
Lecturer: Danica Kragic

EXAM

Image Processing and Computer Vision, DD2422
Tuesday, 10th of February 2009, 14.00–19.00

Allowed helping material: Calculator, the mathematics handbook Beta (or similar).

Language: The answers can be given either in English or Swedish.

General: The examination consists of Part A and Part B. For the passing grade E, you have to answer correctly at least 80% of Part A. If your score is less that 80%, the rest of the exam will not be corrected. Part B of the exam consists of **seven** exercises that can give at most 50 points.

The bonus credits from the labs will be added to Part A if you do not reach 80% - otherwise they will be added to Part B.

The results will be announced within three weeks.

Part A

Provide short answers to the questions! Each answer is worth maximum one point.

1. Why does the epipolar constraint simplify the stereo matching problem?

2. Write a Sobel filter mask and explain what it consists of.

3. Explain what is meant by binomial kernels, give an example of these and give examples of where these are commonly used.

4. Explain terms sampling and quantization.

5. What is "contrast reversal" in terms of gray-level transformations? Draw the corresponding linear transformation.

6. Write filter masks that you used in lab3 to approximate the first order derivative.

7. Write filter masks that you used in lab3 to approximate the second order derivative.

8. What are intrinsic camera parameters?

9. What are vergence and gaze angles (draw a picture)?

10. What is the difference between the perspective and ortographic projection?

11. State briefly how we can detect corners in an image.

12. Give the equation for Bayes theorem and explain each term in the equation

13. How is the Laplacian filter estimated? Give expression and an example two-dimensional mask.

14. What is meant by morphological processing?

15. Describe what scaling, rotation and translation of an image in the spatial domain correspond to in the spectral domain.

16. How can we decide if a two-dimensional filtering kernel/mask is separable?

17. Under perspective projection, a square in the world projects to a quadrilateral in the image. Given the quadrilateral, how can you find the image projection of the center of the square?

18. What is meant by the Laplacian of Gaussian (LoG) filter and how can it be approximated?

19. For what configuration of lines will their vanishing point be in the infinity, that is, it will not be in the image plane?

20. Why do we need to threshold on the gradient magnitude when detecting edges?

Part B

Exercise 1 (3+1+1+3+2=10 points)

Consider a perfect perspective projection camera with focal length 24 mm and a CCD array of size 16 mm×12 mm, containing 500×500 pixels.

Field of View (FOV) is defined as the angle between two points at opposite edges of the image (CCD array), either horizontally or vertically. Thus there are two FOVs, one horizontal and one vertical. Assuming the image center is the center of the image, then FOV is twice the angle between the optical axis and one edge of the image.

1. Give a general expression for computing FOV from focal length and image width.

2. Compute the horizontal FOV and vertical FOV of the given camera.

3. Comment on how FOV affects resolution in an image.

4. Give an expression for computing the pixel coordinates of a point in a 3D scene that is given in camera-frame coordinates. Assume the upper-left corner pixel is (0,0).

5. Compute the pixel where a scene point at coordinates (12m,7m,103m) is imaged.

Exercise 2 (2+2+4+2=10 points)

Your task is to invert a distorted image on the left to the undistorted image on the right by calculating the projective transformation between the two images.

1. Explain how we can see from the left image that it is a result of a projective transformation.

2. Explain what type of a mathematical model can be used to model the undistortion process from the left to the right image. Provide the equation.

3. Write in short a pseudo code and explain how the equation system from above is solved.

4. Are there any particular things we need to be careful with once the model is estimated and the left image is to transformed to the right one?

Exercise 3 (2+1+2=5 points)

Determine for the given binary image:

1. the moments $m_{00}, m_{01}, m_{11}, m_{20}$

2. the center of gravity x_0, y_0

3. the central moments $\mu_{00}, \mu_{01}, \mu_{02}$

$$
\begin{array}{ccccc}
1 & 1 & 1 & 1 & 1 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1 \\
1 & 0 & 0 & 0 & 1 \\
0 & 1 & 1 & 1 & 0 \\
\end{array}
$$

Exercise 4 (4+2+2=8 points)

1. Calculate the PCA-transform to the following sample points:

$$
x_1 = \begin{bmatrix} -2 \\ 0 \end{bmatrix}, \quad x_2 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}, \quad x_3 = \begin{bmatrix} 0 \\ 3 \end{bmatrix}, \quad x_4 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad x_5 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad x_6 = \begin{bmatrix} 2 \\ 4 \end{bmatrix}
$$

2. Transform the sample points to one dimension. What is the mean square error?

3. Calculate the elongatedness of the region that is spanned by the sample points.

Exercise 5 (1+2+2+2+2=9 points)

1. What was the reason for using function **fftshift** in Lab2?

2. Explain the need for the logarithmic intensity transformation in relation to the Fourier transform.

3. What was the difference between functions **gaussffta** and **gaussfftb** in Lab2? What did you use these functions for?

4. Explain how you performed sharpening in Lab2.

5. What were the requirements for detection of edges in Lab3?

Exercise 6 (3 points)

A camera is mounted on the end effector of a robot arm as shown below. The end effector coordinate system E has its Z axis pointing up, its Y axis pointing to the right in the figure, and the X axis pointing out of the page. The origin of the camera C is located at a position of (X,Y,Z)=(0,10,5)cm with respect to E. The X axis of the camera is aligned with the X axis of the end effector. The Z axis of C is tilted down at an angle of 45 degrees with respect to the Y axis of E. Write the 4x4 transformation matrix relating the pose of the camera with respect to the end effector, $^T E_C$.
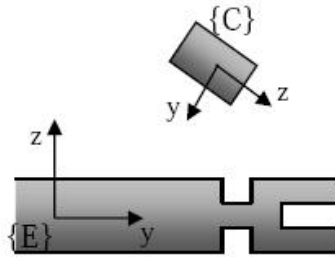
Exercise 7 (3+2=5 points)

1. Give an example of a mask that approximates the derivative in x-direction d_x. Show how higher order derivatives, d_{xx} and d_{xxx} can be estimated using d_x.

2. An image has been smoothed with the following kernel:

$$g = k \cdot [1, 5, 10, 10, 5, 1]$$

Define what is the simplest kernel that we can use to apply multiple convolutions with in order to achieve the same result as above. What should the normalizing factor k be set to?