Computer Science and Communication, KTH
Danica Kragic

Examination

Computer Vision course 2D1421

Monday, 13^{th} of March 2006, 13.00–19.00

Allowed helping material: Calculator, mathematics handbook Beta and a hand-written (not copied) sheet of paper in A4 format with your personal notes. Your notes have to be handed in together with your answers and will be returned to you after the answers are corrected.
Language: The answers can be given either in English or Swedish.
General: The examination consists of **seven** exercises that can give at most 50 credits. To pass the examination you need the half of all credits. The bonus credits will be added to the total sum of your credits given that you passed the laboratory exercises on time during this year (these can give you at most 5 points). The results will be announced within three weeks.

Exercise 1 (5*2=10 credits)

Answer only 5 of the following questions. If you answer more than 5, only the first 5 answers will be corrected and the rest ignored.

(1) In what cases is spectral filtering more appropriate than spatial one? Give two examples.

(2) What is an "ideal low-pass filter"? Is this filter suitable to use in terms of image processing? If yes, give an example of its application. If no, explain why.

(3) Describe basics and draw figures of perspective and orthographic camera models. Given a set of parallel lines in 3D - explain what is the difference in their image projections for both models.

(4) What are the common image point distance measures? Give at least two examples.

(5) Explain the neighborhood concept is terms of digital images. What are the problems related to different concepts?

(6) Describe what scaling, rotation and translation of an image in the spatial domain result in in the spectral domain.

(7) Explain terms sampling and quantization.

(8) What is "contrast reversal" in terms of grey-level transformations? Draw the corresponding linear transformation.

(9) What is the basic idea of an "split-and-merge" algorithm? When is it commonly used?

(10) If a camera views a planar square, what characteristics in the image should be approximately true in order for an affine camera model to be appropriate to use?

Exercise 2 (2+2+2=6 credits)

(a) Given the following perspective camera matrix:

$$P = \begin{bmatrix} 5 & -14 & 2 & 17 \\ -10 & -5 & -10 & 50 \\ 10 & 2 & -11 & 19 \end{bmatrix}$$

and a 3D point in homogeneous coordinates $X = [0\ 4\ 4\ 2]^T$

- What are the Cartesian coordinates of the point X in 3D ?
- What are the Cartesian image coordinates, (u, v) of the projection of X ?

(b) What is a vanishing point and under what conditions will a vanishing point for a perspective camera model be at infinity ? Does a vanishing point exist for an affine camera model and why? Motivate your answer.

(c) What is meant by "camera calibration"? Present it for a perspective camera model and define how many parameters have to be estimated (or in other words: How many degrees of freedom does a perspective camera model have?).

Exercise 3 (2+1+1+1+2+2=9 credits)

(a) A discrete approximation of the second derivative $\frac{\partial^2 I}{\partial x^2}$ can be obtained by convolving an image $I(x, y)$ with the kernel

$$1 \quad -2 \quad 1$$

Use the given kernel to derive a 3×3 kernel that can be used to compute a discrete approximation to the 2D Laplacian. Apply the Laplacian kernel to a center pixel of the following image (Show all the convolution steps!):

$$\begin{matrix} 3 & 2 & 1 \\ 6 & 5 & 4 \\ 9 & 8 & 7 \end{matrix}$$

(b) Why is it important to convolve an image with a Gaussian before convolving with a Laplacian? Motivate your answer by relating to how a Laplacian filter is defined.

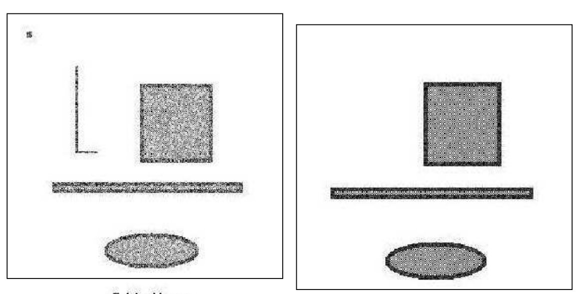
(c) Let us assume doing the following two operations: 1) I first convolve an image with a Gaussian and then take the Laplacian, $\nabla^2(G * I))$, and 2) I first apply the Laplacian to the Gaussian and then convolve the image, $(\nabla^2\ G) * I$. Will the results be the same? If yes - why? If no - why?

(d) Can a Laplacian of a Gaussian (LoG) operator be approximated in any other way than using a sum of second order derivatives? If yes, how?

(e) Given the image below before (left) and after (right) a smoothing filter was applied. The size of the filter is shown as a small square in the upper-left corner in the image (as you can see, its size is rather small compared to the image size). In your opinion, which one of the following filter types most likely produced the image on the right:

1) mean (unweighted averaging) filter,
2) median filter, or
3) Gaussian filter.

Motivate your answer.



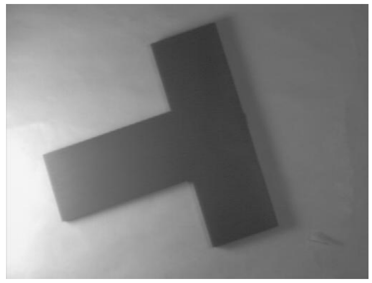
Original image
Smoothing mask size shown in upper left

Smoothed image

(f) Give an example of a mean (unweighted averaging) filter. Is a mean filter in general separable? Why do we prefer separable filters?

Exercise 4 (2+2+2=6 credits)

You are given an image of an object on a background containing a strong illumination gradient. You are supposed to segment the object from the background in order to estimate its moment descriptors.

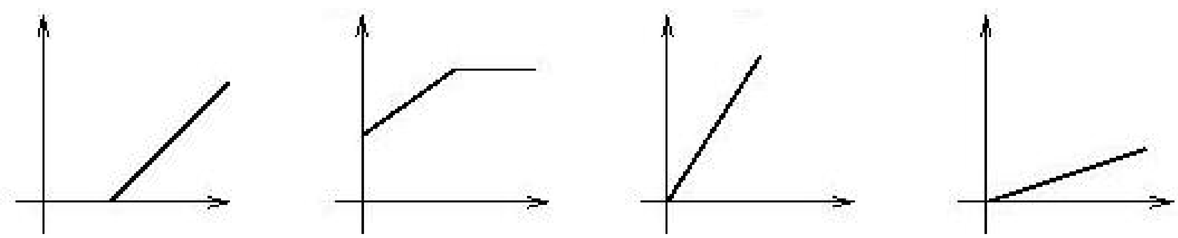


(a) Sketch the histogram of the image and explain what are the problems related to choosing a suitable threshold.

(b) Propose a suitable methodology that could be used to perform successful segmentation.

(c) For the images below, pair them with the correct expressions:

1. Darkening
2. Lightening
3. Compressed to darks
4. Compressed to lights



Exercise 5 (2+2+2=6 credits)

(a) What is the epipolar constraint and how can it be used in stereo matching? Draw a picture!

(b) Assume a pair of parallel cameras with their optical axes perpendicular to the baseline. How do the epipolar lines look like? Where are the epipoles for this type of camera system?

(c) Estimate the essential matrix between two consecutive images for a forward translating camera. What is the equation of the epipolar line for the point p=[x y 1] ?

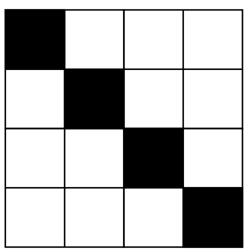
Exercise 6 (2+2+2=6 credits)

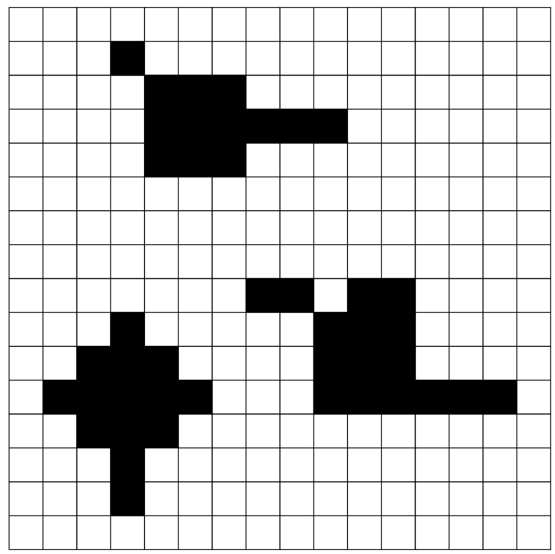
Given the two binary images below.

(a) Create a quad-tree representation (for the image on the left).

(b) Generate a one-dimensional run-length code (for the image on the right) with pixels ordered from top to bottom, left to right.

(c) Perform a morphological opening using a kernel of size 3x3 with all elements equal to one (for the image on the right).





Exercise 7 (2+2+3=7 credits)

- Consider two images taken by a camera that translates right along camera's x-axis direction while viewing a static scene. What does the motion field look like for this motion? Consider the left input image - where is the epipole located? Draw an image of the motion field.
- What is an optical flow constraint equation? Write the equation and explain what each term represents.
- Imagine a robot moving in the environment with a constant velocity. We assume that the floor is perfectly planar and infinitely large. A robot has a camera mounted on the top of the base on height h such that the optical axis of the camera is parallel to the floor. The robot can both rotate and translate on the floor.

  a) Given that you can measure robot's translational and rotational velocity, how would you estimate optical flow for points in the camera image? b) How do the equations change if the robot only translates?