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

EXAM

Image Processing and Computer Vision, DD2422

Tuesday, 11^{th} of February 2008, 14.00–19.00

Allowed helping material: Calculator, the mathematics handbook Beta (or similar) and a handwritten (not scanned or coppied) A4 paper with notes.

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. Write Sobel filter mask and explain what it consists of.
2. How do we define a Laplacian filter mask?
3. What is “fovea” and what is its function?
4. What are “epipolar lines” and why are these important in stereo matching?
5. Mention at least two differences between “cones” and “rods”.
6. Which data reduction technique did we present in the course?
7. In what cases is spectral filtering more appropriate than spatial one? Give two examples.
8. Explain what is meant by binomial kernels, give an example of these and give examples of where these are commonly used.
9. Explain the neighborhood concept is terms of digital images.
10. Describe what scaling, rotation and translation of an image in the spatial domain correspond to in the spectral domain.
11. Explain terms sampling and quantization.
12. What is “contrast reversal” in terms of gray-level transformations? Draw the corresponding linear transformation.
13. What is the basic idea of an “split-and-merge” algorithm?
14. What is meant by “ideal low-pass filter”? Is this filter suitable to use in terms of image processing?
15. What is meant by “blind spot” in humans?
16. Write filter masks that you used in lab3 to approximate the first order derivative.
17. Write filter masks that you used in lab3 to approximate the second order derivative.
18. What are intrinsic camera parameters?
19. What properties are preserved under perspective projection?
20. What are vergence and gaze angles (draw a picture)?

Exercise 1 (2+1+3=6 points)

- Draw images and explain differences between the perspective and orthographic projection. Write the model equations for each of them.
- For what lines will the vanishing point be at the infinity (never project to the image plane)?
- Given a homogeneous transform constructed from rotation matrix R and translation t, derive the form of the inverse.

Exercise 2 (2+2+1+2=7)

- Give an example of a mean (unweighted averaging) filter and explain how we apply a median filter. Are these filters separable?
- You are supposed to find positions of maximum rate of change in an image. Give at least two convolution based methods that can be used for this purpose.
- For typical images, in which part of the frequency spectrum is image noise most severe? For which frequencies will on the other hand image data usually dominate the noise?
- What are the coefficients of the binomial filter of size 5?

Exercise 3 (1+2+2+1=6 points

- What is the epipolar constraint and how can it be used in stereo matching? Draw a picture!
- 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?
- 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 4 (3+3+1=7 points)

Histogram equalization can be used to transform the histogram of an image to the histogram of another. This is beneficial for operations like stereo matching in which image data is matched between two different cameras, that might have very different characteristics.

(a) Assume that the histogram in an image is given by $p(z) = 5z^2 - 3z + 5/6,\ z \in [0,1]$. Determine a transformation $z' = T(z)$, such that the histogram in the new image is $p'(z') = 1,\ z' \in [0,1]$. For which values of z does the transformation result in stretching?

(b) Further, assume we wish to transform an image with the histogram $p(z) = (5-2z)/4,\ z \in [0,1]$ into a new histogram $p'(z') = 2z',\ z' \in [0,1]$. What is the transformation in this case?

(c) Let's say we want to use thresholding to segment a dark cup from a somewhat lighter table. How should the threshold be chosen? Unfortunately, no single threshold results in a satisfactory segmentation. Could you mention (and explain) any other method that could lead to better results?

Exercise 5 (2+4+2=8 points)

- What is meant by classification in computer vision? What features can we use for the purpose if classification?
- For a classification problem with two classes C_A and C_B, are the a priori probabilities $p_A = 1/5$ and $p_B = 4/5$. Assume the following pdf-s

$$p(\bar{z}|C_k) = \frac{1}{2\pi|det\Sigma_k|^{1/2}} e^{-(\bar{z}-m_k)^T \Sigma_k^{-1} (\bar{z}-m_k)/2}$$

  with

$$m_A = m_B = 0,$$

  and

$$\Sigma_A = \begin{pmatrix} 5 & 0 \\ 0 & 1 \end{pmatrix}, \Sigma_B = \begin{pmatrix} 1 & 0 \\ 0 & 6 \end{pmatrix}.$$

  Estimate decision boundaries for the problem.
- Draw an image of the above pdfs and corresponding decision boundaries.

Exercise 6 (2+2+2+2+2=10 points)

- Explain shortly the advantages and disadvantages of functions BINSUBSAMPLE and RAWSUBSAMPLE from Lab1.
- Explain the characteristics of the logarithmic intensity transformation that you used in Lab1 and when it often is a good model to enhance image quality.
- Explain and illustrate with a figure how a position (p, q) in the Fourier domain will be projected in the spatial domain (remember what you did in Lab2).
- Explain how you performed sharpening in Lab2.
- What were the requirements for detection of edges in Lab3?

Exercise 7 (1+2+3=6 points)

This is a noisy binary image of the character L.

- Generate a one-dimensional run-length code (for the image on the right) with pixels ordered from top to bottom, left to right.
- What methodology from the course you would use to remove the noise without too much damage to the character.
- Using the given empty images (as many as you think is needed), explain and perform the operations you think are suitable and draw the result.