Numerisk analys och datalogi, KTH

Mårten Björkman

EXAM

Bildbehandling och datorseende 2D1421
Friday, March 18^{th} 2005, 8.00-13.00

Allowed material: Calculator, mathematics handbook (e.g. Beta) and a hand-written (not copied) sheet of paper in A4 format with your own personal notes. These notes have to be handed in together with your answers and will be returned after answers have been corrected.
Language: Answers can be given in either English or Swedish.
General: The examination consists of **six** exercises that can give at most 50 credits. To pass the examination you need about half of all credits. The bonus credits (at most 5) will be added to the total sum of your credits, given that you passed the laboratory exercises on time during the course of this year. The results will be announced within three weeks.
Course evaluation: We would appreciate if you fill in the evaluation form available on the website.

Exercise 1 (5*2=10 credits)

Answer *five* out of the following *seven* short questions. If you respond to more than *five* questions, only the first *five* will be corrected and counted.

(a) What is the difference between "perspective" and (scaled) "orthographic" projection?

(b) Explain the "cyclopean eye" and why we are usually interested in studying it.

(c) How is a signal in the Fourier domain affected by a translation and scaling in the spatial domain?

(d) What are the three steps of the "Canny" edge detector?

(e) Shortly describe how "K-means" works and what its purpose is.

(f) What is a "Fundamental matrix" and what is it used for?

(g) Mention one common method for "lossy" compression and one for "loss-less" compression.

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

Noise is a common problem is computer vision and careful steps have to be taken to reduce the influence of this noise. Typically, you blur images slightly before you do any further processing. Often we assume the noise to be Gaussian.

(a) 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?

(b) Blurring is not only used to reduce noise. What is the purpose of blurring in scale-space theory?

(c) Assume we reduce the influence of noise by applying a 2D separable binomial filter. What does it mean that a filter is separable and why is seperability a preferable feature?

(d) Given the following portion of an image, apply the 1D binomial kernel $h = \frac{1}{4}[1, 2, 1]$. The results should be 7 pixels in size, i.e. we disregard the boundaries.

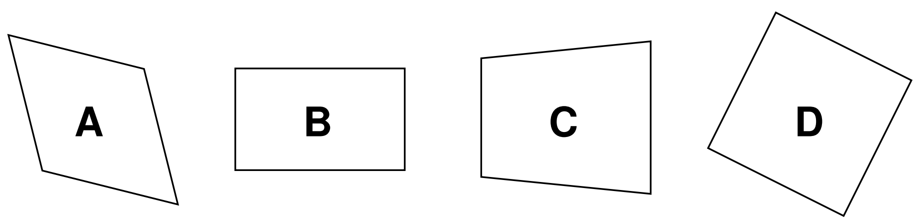
17	18	21	15	13	8	2	1	3

(e) After blurring with the binomial filter in (d) apply a centered differential kernel to the result. Also apply the same differentiation kernel to the original image portion without blurring. How many local maxima (in magnitude) will you get in each case? You only need to consider the center 5 pixels of the results.

Exercise 3 (2+2+2+2=8 credits)

The projection of a two-dimensional surface can be described by a transformation, from the local coordinate system of the surface to the coordinate system of the image plane.

(a) Assume we have a square located somewhere in a 3D scene. Depending on the viewing conditions, the projection of this square might look like the examples below.



Which of these examples can **not** be described by an "affine transformation"? How can you directly see whether an affine transformation is possible?

(b) For at least two of the remaining three examples above, describe how the corresponding affine transformations look like, if we assume that the edges of the square are parallel to the axes of the surface coordinate system and the origin is in the centre. More explicitly, without determining any specific values, what are the relations between the parameters a_{ij} in the affine transformation

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

for each of these two cases?

(c) Assume we have a moving camera and look at two images taken at different points in time. If we have a rigid scene, we can express the "optical flow" for each points in the scene, given that we have a set of motion parameters and the depth of each point. For which motions can we compute the optical flow without knowing these depths? What is the relation between the motion parameters and the optical flow for points located at infinity?

(d) The so called Optical Flow (or Brightness Constancy) constraint is given by $L_x u + L_y v + L_t = 0$. How can this constraint be derived? How does the “aperture problem” affect the application of the constraint?

Exercise 4 (3+3+3=9 credits)

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+3+1=6 credits)

In object recognition and classification we often like to go from a high-dimensional image space to a lower-dimensional feature space, in which comparisons between new images and previously stored models can be made.

(a) In order for comparisons to be successful, they have to be made invariant to a number of changes that frequently appear in real images. Mention at least three such invariances.

(b) One way of reducing the dimensionality is by applying Principle Component Analysis (PCA). From a set of training images $\{\mathbf{X_i}\}$ we compute a number of eigenimages (eigenspaces, eigenfaces) $\{\mathbf{U_k}\}$, that represent a basis with which every (training and test) image can be described.

- Given the eigenimages, each new image $\mathbf{X}$ can be represented by a set of coefficients $\{c_k\}$. How do you compute these coefficients?
- What will the number of coefficients you need depend on?
- Once you have the coefficients for a particular image, how do you compare these to those of other images (that are usually stored in a database)?

(c) If you compare PCA representations the way you described in (b), will the invariances you mentioned in (a) be satisfied?

Exercise 6 (3+2+3=8 credits)

Using Taylor expansion

$$f(x+h) = f(x) + hf'(x) + \frac{h^2}{2!}f''(x) + \frac{h^3}{3!}f'''(x) + O(h^4),$$

where h is the distance between two neighbouring pixels, it can be shown that a result of a differentiation kernel $d_x = \frac{1}{2}[-1,0,1]$ applied to an image $f(x)$ is

$$d_x * f(x) = \frac{1}{2}(f(x+h) - f(x-h)) = hf'(x) + \frac{h^3}{6}f'''(x) + O(h^4)$$

Thus, a derivative approximation based on d_x will result in a dominating error proportional to h^3 (if we assume that $h \ll 1$). Unfortunately, we cannot find a better approximation of a first order derivative than d_x, if the kernel is limited to only 3 pixels in width.

(a) Using the same method based on Taylor expansion, determine how well $d_{xx} = [1,-2,1]$ approximates a second order derivative.

(b) One might wonder whether a better approximation of a first order derivative is possible, if we increase the length of the kernel. Without knowing the best kernel, can you tell whether it ought to be symmetric or anti-symmetric? How many parameters do we have to determine if we search for a kernel of length 5?

(c) Find a differentiation kernel of length 5, such that also the error term proportional to h^3 disappears. Which error term will now dominate?

Good luck!