Numerisk analys och datalogi, KTH
Mårten Björkman

EXAM

Bildbehandling och datorseende 2D1421
Wednesday, March 9th 2005, 14.00-19.00

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 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) In what sense is vision an “active process”?

(b) Mention at least two differences between “cones” and “rods”.

(c) What is a “neighbour” and what is a “connected component”? Show with an illustration.

(d) What is an “ideal” low-pass filter and why is such a filter not suitable for computer vision?

(e) What is the difference between “optical flow” and “motion field”?

(f) What are “epipolar lines” and why are these important in stereo vision?

(g) What is “entropy” and why is it of interest in image compression?

Exercise 2 (5+2=7 credits)

We have a set of five 2D points; $\mathbf{p_1} = (-2,-1)$, $\mathbf{p_2} = (-1,-2)$, $\mathbf{p_3} = (0,0)$, $\mathbf{p_4} = (1,1)$ and $\mathbf{p_5} = (2,1)$.

(a) Assume the points originate from a two-dimensional distribution. Based on the few points given, characterize this distribution by an ellipse. Compute and draw this ellipse.

(b) Alternatively, assume that the points come from a one-dimensional distribution, i.e. they are placed along a line. Find this line and compute the mean square distance of the points to the line.

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

Common for many operations in computer vision is the reduction of data. This is true for both object recognition and image compression.

(a) Why do we like (or need) to reduce data in object recognition and image compression respectively? What is the difference in the data we ignore?

(b) What kind of redundancies can be exploited in image compression? Mention at least two kinds of redundancies and explain them.

(c) Matching of features for recognition has to be made invariant to, among other things, 1) scale, 2) illumination and 3) rotation. Describe how this can be done in practice. Mention at least one typical example for each kind.

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

Assume we have a camera defined by a perspective projection, $\mathbf{x} = \mathbf{MX}$, with the projection matrix

$$\mathbf{M} = \begin{pmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{pmatrix}$$

and the homogeneous world and image coordinates given by $\mathbf{X} = (X,Y,Z,1)^T$ and $\mathbf{x} = (x,y,1)^T$.

(a) The matrix $\mathbf{M}$ can be expressed in terms of intrinsic and extrinsic camera parameters. Give at least two parameters of each kind.

(b) How would $\mathbf{M}$ look like if it were an "affine" projection matrix? Under what assumptions is an affine projection feasible?

(c) Assume we have a plane in 3D space defined by the equation

$$2X - Z + 2 = 0.$$

Introduce two coordinate axes in this plane, so that a point on the plane is given by the coordinates $\mathbf{x}' = (x', y', 1)^T$. Show that the projection from the 3D plane to the image plane can be expressed as $\mathbf{x} = \mathbf{A}\mathbf{x}'$, where A is a 3×3 projective transformation. What is A for your choice of coordinate axes? Hint: Use for example $(1,0,2,0)^T$ and $(0,1,0,0)^T$ as axes with $(0,0,2,1)^T$ as origin.

Exercise 5 (2+2+5=9 credits)

Assume we have a robot with cameras placed $h = 120$ cm above an ugly coloured plastic floor bought by your grandmother in the late 60s. The texture of the floor is orange with lots of large brown circles on.

(a) The cameras are tilted towards the floor such that a particular circle is projected in the center of the left camera. Measured in the image the projection is $w_c = 40$ pixels in width and $h_c = 15$ pixels in height. How far away from the left camera is the circle on the floor?

(b) Assume that the two cameras are placed perfectly in parallel and have the same intrinsic parameters, with focal lengths equal to $f = 600$ pixels. How large is the baseline between the cameras, if the disparity in the center of the projected circle is $d = 24$ pixels? What is the diameter of the circle measured in centimetres?

(c) Grandmother's walls are neatly coloured blue. To separate images into wall and floor pixels we apply pixel classification, using a single blue-yellow colour channel, $z \in [\,0, 255\,]$. We assume the classes of pixels to be described by normal distributions

$$p(z \mid C_k) = \frac{1}{\sqrt{2\pi}\sigma_k} e^{-(z-m_k)^2/(2\sigma_k^2)},$$

with standard deviations $\sigma_W = 80$ and $\sigma_F = 40$, and means $m_W = 60$ and $m_F = 180$. If the prior probabilities are $p(C_F) = 0.7$ and $p(C_W) = 0.3$, between which colour values will a pixel z be classified as a floor pixel?

Exercise 6 (3+4+2=9 credits)

The primary reason for expressing discrete filters in frequency space is to understand their behaviours, in particular in relation to their equivalents in continuous space.

(a) Show using the definitions of convolutions and Fourier transforms that the Fourier transform of a convolution of two kernels is the same as the product of the Fourier transforms of each kernel, i.e. $\mathcal{F}(h * g) = \mathcal{F}(h)\mathcal{F}(g)$.

(b) The Fourier transform of a continuous second order derivative is $\mathcal{F}(\delta_x^2) = -\omega^2$. Unfortunately, on discrete data we have to approximate these derivatives. Assume that we twice apply a first order differentiation kernel $h_x = \frac{1}{2}[-1, 0, 1]$. What is the frequency response of this approximation? What is the frequency response if we instead apply the second order differentiation $h_{xx} = [-1, 2, -1]$?

(c) Draw the frequency responses of the approximations in (b) as functions of ω. Which alternative is preferable, if we are looking for an approximation to a continuous second order derivative?

Note: If you like to know whether you passed or failed this exam before the next exam, which is given on **March 18th, 08.00-13.00**, write your email address on top of the hand-ins.

Good luck!