Instructions TEN1

The exam consists of 4 parts, each corresponding to one of the four lab exercises of TSBB12 / TS1017.

- Part 1 covers tracking
- Part 2 covers motion
- Part 3 covers denoising
- Part 4 covers stereo

Each part contains 3 tasks, two that require description of terms, phenomena, relations, etc. (type A) and one that goes more into detail and may require some calculations (type B).

Correct answers for type A give 2p and for type B give 4p, i.e., each part gives 8p, and a total of 32p for the whole exam.

In order to pass with grade 3, at least 15p are required.

In order to pass with grade 4, at least 22p are required.

In order to pass with grade 5, at least 27p are required.

TSBB12: If you have less than 15p, but in only one or two of the four parts you have less than 4p, you can apply for complementing the exam at the next ordinary examination occasion. The application is subject to acceptance by the examiner. If the complementing part is passed, the whole exam is passed with grade 3 and dated the day of the complementing exam.

TS1017: As this is a midterm examination (on Parts 1 and 2) and not a formal examination, the following rule applies: If you have 8p or more in Parts 1 and 2, you can choose to use this result for the formal examination of Parts 1 and 2, i.e., you can choose to add the points from this midterm examination of Parts 1 and 2 to the achieved points on Parts 3 and 4 in the formal examination.

All tasks should be answered on **separate sheets** that are to be attached to the exam.

Write your AID-number and the date on all paper sheets that you attach to the examination. In addition, these sheets should be numbered in consecutive order.

Good luck! Michael Felsberg, Klas Nordberg, and Per-Erik Forssén

PART 1: Tracking

Task 1 (A, 2p) A stationary surveillance camera is imaging a section of an indoors corridor. During the day, the corridor is illuminated by approximately constant natural light. The ceiling also has fluorescent lights, and these can be switched on and off.

a) Write down a mixture distribution $p(I)$ for the intensity of a pixel on a uniformly coloured wall in the scene. Explain, and name all entities in your expression.
b) Lets say that the mixture components above are Gaussians. Explain the meaning of the mean and the standard deviation of each such Gaussian.

Task 2 (A, 2p) Consider KLT tracking of a 1D signal. The KLT tracker is based on the Taylor expansion of the shifted image $f(x+d)$. Now, instead make a Taylor expansion of the blurred image $(f*g)(x)$, where $g(x)$ is a Gaussian filter.

a) One of the terms in the expansion has the form $\frac{\partial}{\partial x}(f*g)(x)$. Derive the filter that can be used to compute these derivatives.
b) The solution for the displacement can be written $d = (I(x) - J(x))/G(x)$, where $G(x)$ is the output of the filter you just derived, but what are I and J?

Task 3 (B, 4p) *Symmetric* Lucas-Kanade (KLT) tracking measures the displacement vector $\mathbf{d}$ between two image regions f and g by splitting the displacement into two halves. This is achieved by minimizing the error

$$\varepsilon(\mathbf{d}) = \iint (f(\mathbf{x} + \frac{1}{2}\mathbf{d}) - g(\mathbf{x} - \frac{1}{2}\mathbf{d}))^2 w(\mathbf{x})\, d\mathbf{x} \ .$$

Derive an expression for how $\mathbf{d}$ can be estimated from f, g.

PART 2: Motion

Task 4 (A, 2p) Local displacements of regions in an image sequence can be estimated, for example, using block matching or based on the Lucas-Kanade equation over multiple scales, or a combination of the two methods. Generate a table with characteristics of these two approaches where they differ. Discuss why and how they can be combined in a suitable way.

Task 5 (A, 2p) Motion analysis is often based on the brightness constancy constraint (BCCE): $\frac{\partial f}{\partial x_1}v_1 + \frac{\partial f}{\partial x_2}v_2 + \frac{\partial f}{\partial t} = 0$. Now, consider the situation where $f(x_1, x_2) = ax_1 + bx_2 + c$, where $a, b, c > 0$, and $\frac{\partial f}{\partial t} = d$. We shall now try to find out what BCCE tells us about $\mathbf{v} = (v_1, v_2)^T$!

a) Express BCCE in terms of a, b, c, d! If you plot all the valid solutions for $\mathbf{v}$ in the image plane, what do you get?
b) Which is the shortest velocity vector that satisfies BCCE? (This is called the normal velocity.)

Task 6 (B, 4p) A local velocity estimate (v_x, v_y) at point (x, y) and time t can be determined by solving the BCCE based on data in a local region around the point. This can be done by either collecting data in a 2D region around the point, leading to an equation based on the 2D structure tensor, or collecting data in a 3D region, leading to an equation based on the 3D structure tensor.

a) What types of equations do you get in each of the two cases? How is (v_x, v_y) related to the structure tensor in each of the two cases?
b) The two equations have a qualitative difference in terms of how they deal with the aperture problem. Describe this difference?

PART 3: Denoising

Task 7 (A, 2p) The diffusion equation in the case of anisotropic diffusion can be written

$$\frac{\partial}{\partial s} L = \frac{1}{2}\,\mathrm{div}(\mathbf{D}\,\mathrm{grad}\,L)$$

where $\mathbf{D}$ is the local diffusion tensor. A useful approximation of this equation is

$$\frac{\partial}{\partial s} L \approx \frac{1}{2}\,\mathrm{trace}(\mathbf{D}\,\mathbf{H}L)$$

where $\mathbf{H}L$ is the local Hessian (second order derivatives in x and y) of L. What critical assumption is this approximation based on, and why is the assumption reasonable?

Task 8 (A, 2p) In image enhancement, the local orientation tensor $\mathbf{T}(\mathbf{x})$ is mapped to a control tensor $\mathbf{C}(\mathbf{x})$ and it is the latter that controls the locally adapted filter at point $\mathbf{x}$. Describe how this mapping is done, i.e., in the case of enhancement of 2D images.

Task 9 (B, 4p) The Perona-Malik diffusion can be derived by minimizing the functional

$$\varepsilon(f) = \frac{k^2}{2}\int_{\Omega} \exp\left(-\frac{|\nabla f|^2}{k^2}\right)\,dx\,dy$$

where k is a constant. What is the Lagrangian $L(f, f_{x_w})$ of this functional? What is the Euler-Lagrange equation $L_f - \sum_w \partial_{x_w} L_{f_{x_w}} = 0$ for this functional? What is the resulting evolution equation (hint: apply gradient descent)? The result may be presented neglecting constant factors.

PART 4: Stereo

Task 10 (A, 2p) Given a set of points in each of two stereo images, the RANSAC algorithm is used to determine pairs of corresponding points. Describe the computational steps are made for each iteration of the RANSAC loop.

Task 11 (A, 2p) Local invariant features are usually characterised by their *geometric invariances* and *photometric invariances*. List two geometric invariances, and two photometric invariances, and explain what they are useful for.

Task 12 (B, 4p) Assume that we have three images of the same scene, labeled 1, 2, and 3, taken with pinhole cameras that have distinct camera centers. A large set of interest points is visible in all three images, and can be used to estimate three fundamental matrices: $\mathbf{F}_{12}, \mathbf{F}_{23}, \mathbf{F}_{31}$, where $\mathbf{F}_{ij}$ is the fundamental matrix between images i and j. From each $\mathbf{F}_{ij}$ we can determine the epipolar points $\mathbf{e}_{ij}$, each being the image of camera center j in image i.

a) Each of the 6 epipolar points must satisfy a simple relation with one of the three fundamental matrices. Describe these 6 relations.

b) In addition, each of the 3 fundamental matrices must satisfy a relation together with pairs of epipolar points. These relations do not follow from the first set of relations. Describe these 3 relations.