Information page for written examinations at Linköping University

Examination date	2010-03-10
Room (1) If the exam is given in different rooms you have to attach an information paper for each room and <u>mark intended place</u>	T1
Time	8-12
Course code	TSBB12
Exam code	TEN1
Course name **Exam name**	Datorseende teori Skriftlig tentamen
Department	ISY
Number of questions in the examination	12
Teacher responsible/contact person during the exam time	Michael Felsberg
Contact number during the exam time	013 282460
Visit to the examination room approx.	10:00
Name and contact details to the course examinator (name + phone nr + mail)	Michael Felsberg, mfe@isy.liu.se
Equipment permitted	Dictionary Swedish-English-Swedish
Other important information	See instructions on the next page

Instructions TEN1

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

- 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, phenomenon, 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.

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.

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 simple way to detect and track moving objects is to use a mixture model of the background. Write down a mixture model of pixel intensity. Name the algorithm that estimates a mixture model from measurements, and describe how it works.

Task 2 (A, 2p) The KLT tracker makes use of a continuous interpolation $f(x, y)$ of the discrete image $b_{m,n}$, and its x- and y- derivatives f_x and f_y to find displacements with sub-pixel accuracy. For best results, the algorithm should use regularised derivatives. Explain why.

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

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

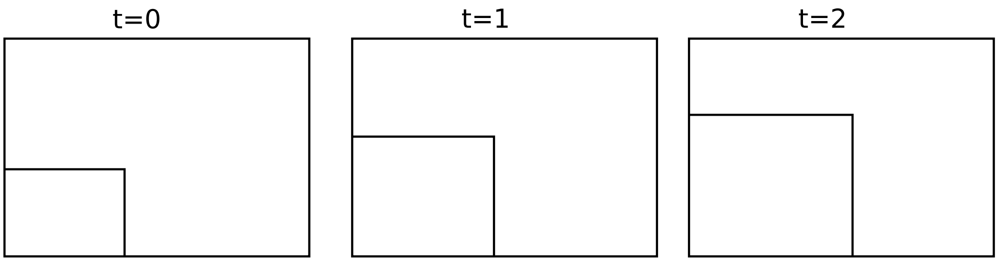
Derive an expression for how $\mathbf{d}$ can be estimated from f, g.

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 in image sequences is sometimes based on the assumption that the 3D motion vectors of 3D points in the scene, when projected onto the image plane, generates a *motion field* which can be determined. This assumption is not correct in general. Give two examples of why we cannot determine the motion field from measurements on the image sequence.

Task 6 (B, 4p) An image sequence contains a corner moving towards the top right corner of the image, see the figure below.



- Motion estimation is made on the sequence, based both on local analysis, e.g., using Kanade-Lucas' method, and on global analysis, e.g., based on Horn-Schunck's method. Sketch the resulting motion fields for $t = 1$ on a separate sheet.
- Explain possible differences in the resulting motion fields, in particular close to the moving edges and in homogeneous parts of the image.

PART 3: Denoising

Task 7 (A, 2p) Denoising based on anisotropic diffusion can be described as applying a space varying filter onto the image signal. What is the general shape of this filter, which parameters of the filter shape can vary over the image, and what is the qualitative dependency of these parameters on the local structure of the image?

Task 8 (A, 2p) Image enhancement on 2D images can be done by first computing local orientation estimates, e.g., in terms of the structure tensor $\mathbf{T}(\mathbf{x})$. This tensor is then mapped into a control tensor $\mathbf{C}(\mathbf{x})$. Describe how many filters are needed for the final enhancement filtering, what are their general shape in the Fourier domain, and how the enhanced image is computed at a point $\mathbf{x}$, based on the responses of the various filters and $\mathbf{C}(\mathbf{x})$.

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) The RANSAC algorithm can be used to simultaneously determine point correspondences and a fundamental matrix related to a stereo image pair. Describe the computations that take place for each iteration of the RANSAC loop in this case.

Task 11 (A, 2p) In wide-baseline stereo, image patches sampled in canonical frames are used to find correspondences between two images. The frames are typically found using either the SIFT, or the MSER detector. The SIFT frame has geometric invariance to similarities, and MSER has affine geometric invariance. Both of these are special cases of a homography. Express both affine and similarity transformations with homographies. Also write down how many degrees of freedom each of the transformations has.

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.

Assuming that the image coordinates of the interest points contain some noise (which they always do in practice) these three estimated fundamental matrices are not exactly correct relative to the cameras. Furthermore, because of the noise they may not be consistent, i.e., it may not be possible to determine camera matrices $\mathbf{C}_1, \mathbf{C}_2, \mathbf{C}_3$ such that the three fundamental matrices are computed from the camera matrices. Describe a necessary condition on the fundamental matrices such that they become consistent.