# Information page for written examinations at Linköping University

| | |
|---|---|
| **Examination date** | 2012-06-02 |
| **Room** | TER3 |
| **Time** | 14-18 |
| **Course code** | TSBB15 |
| **Exam code** | TEN1 |
| **Course name**<br>**Exam name** | Computer Vision, written examination |
| **Department** | Dept of EE (ISY) |
| **Number of questions in the examination** | 12 |
| **Teacher responsible/contact person during the exam time** | Per-Erik Forssén |
| **Contact number during the exam time** | 013-285654 |
| **Visit to the examination room approx.** | 15.30 |
| **Name and contact details of the course administrator** | Klas Nordberg<br>281634<br>klas@isy.liu.se |
| **Equipment permitted** | Dictionary Swedish-English-Swedish |
| **Other important information** | See instructions on the next page |
| **Which type of paper should be used, cross-ruled or lined** | Cross-ruled |
| **Number of exams in the bag** | 22 |

# Instructions TEN1

The exam consists of 4 parts, each corresponding to one of the four lab exercises of TSBB15.

- Part 1 covers tracking
- Part 2 covers motion
- Part 3 covers denoising
- Part 4 covers multiple view geometry

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.

Re-use of the midterm examination result: If you have written the midterm examination and obtained 8p or more in Parts 1 and 2, you may re-use this result, and not answer the corresponding parts in this 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! Per-Erik Forssén, Michael Felsberg, and Klas Nordberg

# PART 1: Tracking

**Task 1** (A, 2p) Consider using the EM algorithm to estimate the parameters of the following mixture distribution:

$$p(I) = p(I|\Gamma_1)P(\Gamma_1) + p(I|\Gamma_2)P(\Gamma_2)\,,$$

using data $I_1 \dots I_N$, and mixture components of the form $p(I|\Gamma_k) = \frac{1}{\sqrt{2\pi}\sigma_k} e^{-0.5(I-\mu_k)^2/\sigma_k^2}$.
a) Describe how to compute the partial memberships in the E step. (1p)
b) Describe how to update the model parameters in the M step. (1p)

**Task 2** (A, 2p) In KLT tracking, the local displacement $\mathbf{d}$ between two images is found as:

$$\mathbf{d} = \mathbf{M}^{-1}\mathbf{e}\,,\ \text{where}\ \ \mathbf{M} = \sum_{\mathbf{x}\in\mathcal{N}} w(\mathbf{x})\nabla a(\mathbf{x})\nabla a(\mathbf{x})^T\,,\ \mathbf{e} = \sum_{\mathbf{x}\in\mathcal{N}} w(\mathbf{x})\nabla a(\mathbf{x})(b(\mathbf{x}) - a(\mathbf{x}))$$

For the following quantities: $w$, $\nabla a$, $b$. Describe what each of them denotes, (1p)
and explain how they can be determined. (1p)

**Task 3** (B, 4p) In standard Kanade-Lucas tracking, based on translation only, we minimize the cost function:

$$\varepsilon = \sum_{\mathbf{x}\in\mathcal{N}} ||f_1(\mathbf{x}) - f_2(\mathbf{x}+\mathbf{d})||^2$$

By using a first order Taylor expansion, setting the derivative to zero, and solving for $\mathbf{d}$, we obtain the expression given in Task 2 (with a different notation). Now, instead formulate a cost function where the translation is distributed symmetrically in $f_1$ and $f_2$. (1p)
Derive an expression for $d$ from your cost function, in the 1D case, where $x$ and $d$ are scalars. (3p)

# PART 2: Motion

**Task 4** (A, 2p) A common extension to motion estimation is coarse-to-fine estimation.
a) Explain why this is done. (1p)
b) The images $I$, and $J$, at two different resolutions, correspond as $J(\mathbf{x}) = (g * I)(2.0\mathbf{x} - 0.5)$, where $g$ is a Gaussian blur kernel. If we want the flow vector $\mathbf{d}$ at position $\mathbf{x}$ in the resolution of image $I$, where in resolution $J$ should we interpolate to find it, and how should we scale the result? (1p)

**Task 5** (B, 4p) One way to determine optical flow is to formulate a cost function for a *local* region $\Omega$ as

$$\epsilon = \int_\Omega w(\mathbf{x})\|\nabla_3 f \cdot \mathbf{v}\|^2 d\mathbf{x}, \qquad \text{where} \quad \nabla_3 f = \begin{pmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} & \frac{\partial f}{\partial t} \end{pmatrix}^T \qquad \text{and} \quad \mathbf{v} \sim \begin{pmatrix} v_1 & v_2 & 1 \end{pmatrix}^T,$$

and where $\sim$ denotes equality up to scale. We assume that $\nabla_3 f$ is known at all points. We can estimate $(v_1, v_2)$ in two distinct ways that differ in what additional constraint is applied to the spatio-temporal motion vector $\mathbf{v}$. Describe the two constraints, and describe the principal equations that need to be solved in order to obtain $(v_1, v_2)$ in both cases. Compare the two approaches in terms of numerical stability for the case of moving edges/lines as well as in terms of computational complexity.

**Task 6** (A, 2p) Horn & Schunck estimate the motion field by adding an extra term to the error function and solving the motion field *globally*.
a) Write down the cost function, including the extra term, and state the underlying assumption that motivates the term. (1p)
b) Describe some plausible case when this assumption is not valid. (1p)

# PART 3: Denoising

**Task 7** (A, 2p) Linear diffusion corresponds to convolving the image with a Gaussian kernel of suitable width. Non-linear diffusion methods do not have a corresponding shift-invariant operator, but instead correspond to a shift-variant kernel that depends on the local image structure. Please sketch what the support of this kernel looks like for the two situations of (1) (combed) hair at very close distance (individual hairs are visible) and (2) larger distance (hair looks like a noisy structure), both for (a) Perona-Malik diffusion and (b) tensor driven diffusion (4 sketches in total). What can you say about the resulting noise-level after denoising?

**Task 8** (A, 2p) Given a noisy image $f$, an enhanced image $h$ can be computed as

$$h(\mathbf{x}) = \int f(\mathbf{x} - \mathbf{y})\, g_{\mathbf{x}}(\mathbf{y})\, d\mathbf{y}$$

where $g_{\mathbf{x}}$ is filter that is locally adapted to the image structure at position $\mathbf{x}$. Since $g_{\mathbf{x}}$ depends on $\mathbf{x}$, this last expression is not a convolution. Assume that the local image structure at position $\mathbf{x}$ is encoded in some position dependent coefficients $c_k(\mathbf{x})$.
(a) How can $g_{\mathbf{x}}$ be formulated to make it suitable for an effective computation of $h$? (1p)
(b) Given (a), how can $h$ be computed in terms of convolutions and linear combinations? (1p)

**Task 9** (B, 4p) The absolute error is less sensitive to outliers than the quadratic error, which is why regularization using the L1-norm is often considered to be preferable to quadratic regularization.

a) Write down the L1-regularization term as it is commonly used in variational methods for denoising. (1p)
b) Write down the corresponding term of the Euler-Lagrange equation (assume a diffusion process, i.e., no data term is required). (1p)
c) Which case is problematic when calculating the term in b) and how can this problem be avoided? (1p)
d) Give an intuitive interpretation of the term from b) for the case of a 1D signal. (1p)

# PART 4: Multiple View Geometry

**Task 10** (A, 2p) a) Assume we have a method that generates tentative correspondences with each having a probability $p$ of being an outlier. What is the probability of picking 4 inlier correspondences? (1p)
b) Assume we have two image point sets $\mathcal{A}$ and $\mathcal{B}$, both with projections of the same 100 3D points, but from two different views. What is the probability of picking a correct correspondence by drawing one sample from each set randomly? (1p)

**Task 11** (A, 2p) a) Explain how rotation invariance is obtained for the SIFT detector. (1p)
b) Explain how affine illumination invariance is obtained for the SIFT descriptor. (1p)

**Task 12** (B, 4p)

a) Formulate the P$n$P-problem. Describe what are the known entities and what are the unknowns to be solved. (1p)
b) What is the minimum $n$ for the P$n$P problem? (1p)
c) Describe how P$n$P can be used in a 3D-reconstruction pipeline, and explain how P$n$P differs from the bundle adjustment problem. (2p)