102470 - Computer Vision Course
Institut für Informatik
Universität Bern

# MOCK EXAM

21/12/2021

- **You can use one A4 sized hand-written sheet of paper.**
- **No books, notes, computers, calculators and cellular phones are allowed.**
- **The number of points of the exam is 50.**

**Name:** ____________________ **Student ID:** ____________________

| Exercise | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| Total | 10 | 10 | 11 | 19 |
| Mark | | | | |

# Multiple-Choice Questions [10 Points]

**Correct answer: +1 Point, Wrong answer: -1 Point, No answer: 0 Points. Negative total points will be elevated to 0.**

1. **True** **False** The depth of field is a function of the aperture.
   **Solution.**
   True.
   Source: Slides 17-18 of lec02a_camera.pdf
2. **True** **False** If the rows of a 2D convolution filter are linearly independent, the filter is separable.
   **Solution.**
   False.
   Source: Slides 27-28 of lec02b_filter.pdf
3. **True** **False** The gradient of the total variation TV($u$) of a signal $u$ is linear in $u$.
   **Solution.**
   False.
   Source: Slide 86 of lec03-04_energy minimization.pdf
4. **True** **False** A Lambertian surface has the property that the reflected color does not change with the light direction.
   **Solution.**
   False. It is the viewing direction that does not change the color.
   Source: Slide 7 of lec04_shading_photometric_stereo.pdf
5. **True** **False** Shape from shading uses a fixed camera and a still scene illuminated from a specific direction.
   **Solution.**
   True.
   Source: Slide 32 of lec04_shading_photometric_stereo.pdf
6. **True** **False** The only assumption in estimating optical flow is the brightness constancy.
   **Solution.**
   False.
   Source: Slide 9 of lec05a_optical_flow.pdf
7. **True** **False** A homography maps parallelograms to parallelograms.
   **Solution.**
   False. It maps quadrilaterals to quadrilaterals.
   Source: Slide 10 of lec05b_registration.pdf
8. **True** **False** Structure from motion can recover the absolute scale of the scene.
   **Solution.**
   False.
   Source: Slides 4-5 of lec09b_sfm.pdf
9. **True** **False** The SIFT feature extractor (not the descriptor) is invariant to the rotation of the image.
   **Solution.**
   True. Each gradient orientation is relative to the keypoints orientation.
   Source: Slide 43 0f lec06_interest_points.pdf

10. **True** **False** The epipolar geometry depends only on the intrinsic and extrinsic parameters of the cameras.

    **Solution.**
    True.
    Source: Slides 24 and 27 of lec08a_epipolar.pdf

# Photometry, Features, Filters & Photometric Stereo [11 points total]

**[2 points]**

1. Circle the factors below that affect the intrinsic parameters of a camera model.
   (a) Offset of the optical center
   (b) Camera orientation
   (c) Image resolution
   (d) Focal length

   **Solution.**
   a, c, d. [**2 points**]

   Half a point for every correct circle/non circle.

**[5 points]**

2. Suppose that the normal map $\mathbf{n}$ of a depth map $d$ is given. Write the equation that relates the depth map to the normal map.

   *Hint: Write the depth map as a 3D surface and relate it to its tangent vectors.*

   **Solution** If we write the 3D point on the surface corresponding to a pixel $(x, y)$ as [**1 points**]

$$P = \begin{bmatrix} x \\ y \\ d(x,y) \end{bmatrix} \tag{1}$$

then we can write the tangent vectors by taking the first order derivatives with respect to $x$ and $y$ [**2 points**]

$$T_x = \begin{bmatrix} 1 \\ 0 \\ d_x(x,y) \end{bmatrix} \qquad T_y = \begin{bmatrix} 0 \\ 1 \\ d_y(x,y) \end{bmatrix}. \tag{2}$$

Since the normal to the surface must be orthogonal to both tangent vectors, we have [**1 points**]

$$\mathbf{n} = \begin{bmatrix} n_1 \\ n_2 \\ n_3 \end{bmatrix} \propto \begin{bmatrix} \nabla d \\ -1 \end{bmatrix} = \begin{bmatrix} d_x(x,y) \\ d_y(x,y) \\ -1 \end{bmatrix}. \tag{3}$$

we can write $\mathbf{n} = \alpha \begin{bmatrix} d_x(x,y) \\ d_y(x,y) \\ -1 \end{bmatrix}$ and from here we find $\alpha = -n_3$

Then we can write [**1 points**]

$$d_x = D_x d = -\frac{n_1}{n_3} \tag{4}$$

$$d_y = D_y d = -\frac{n_2}{n_3} \tag{5}$$

where $D_x$ and $D_y$ are the matrices denoting derivatives with respect to $x$ and $y$. Thus we have a linear system

$$\begin{bmatrix} D_x \\ D_y \end{bmatrix} d = -\begin{bmatrix} \frac{n_1}{n_3} \\ \frac{n_2}{n_3} \end{bmatrix}. \tag{6}$$

[**1 points**]

**[3 points]** 3. The median filter takes the middle element after sorting the elements in a patch. While this operation is quite robust to non Gaussian noise, it can lead to severe flattening of the texture when the size of the patch is large. How could we modify the median filter so that it retains the original texture while removing outliers?

**Solution** One possible modification is to leave the center pixel if this is within a certain distance from the median pixel of the patch (*e.g.*, below the first quartile or above the third quartile).

# Optical Flow [11 points total]

1. The *brightness constancy* constraint for optical flow states that for all $x$, $y$ and $t > 0$ and some $\Delta t$

$$I(x, y, t) = I(x + \Delta x, y + \Delta y, t + \Delta t). \tag{7}$$

By assuming that $\Delta t$ is small and by using the Taylor approximation, this can be reformulated as

$$\frac{\partial I(x, y, t)}{\partial x}\Delta x + \frac{\partial I(x, y, t)}{\partial y}\Delta y + \frac{\partial I(x, y, t)}{\partial t}\Delta t = 0. \tag{8}$$

**[3 points]**

(a) Describe the meaning of the brightness constancy and all symbols involved in the approximation equation (8).

**Solution.**

The brightness constancy constraint states that the image intensity $I$ of an object moving from pixel location $x$ to location $x + \Delta x$ does not change color over a small time period $\Delta t$. **1 p**

Definition of symbols:

- $\frac{\partial I}{\partial x}$, $\frac{\partial I}{\partial x}$ are the x- and y component of the image gradient. $\frac{\partial I}{\partial t}$ is the change in image brightness over time. **1 p**
- $\Delta x$, $\Delta y$ is the absolute change in pixels in x- and y direction. $\Delta t$ is the time between the two measurements. **1 p**

**[2 points]**

(b) Transform eq. (8) into a form that contains the x- and y component $(v_x, v_y)$ of the optical flow.
*Hint: Recall that the optical flow relates to the velocity of moving points.*

**Solution.**

We divide both sides of eq. 8 by $\Delta t$, yielding

$$\frac{\partial I}{\partial x}\frac{\Delta x}{\Delta t} + \frac{\partial I}{\partial y}\frac{\Delta y}{\Delta t} + \frac{\partial I}{\partial t} = 0. \tag{9}$$

In this form, we have $v_x = \frac{\Delta x}{\Delta t}$ and $v_x = \frac{\Delta y}{\Delta t}$, which are the components of the optical flow.

**[6 points]**

(c) Give three scenarios where optical flow fails. Justify your answer by referring to the formulas above.

**Solution.**

**1.** The motion is considerably larger than 1 pixel. In this case the Taylor approximation is not accurate. **2 p**

**2.** The motion is not uniform locally. **2 p**

**3.** Brightness changes rapidly. The brightness constancy assumption (and the equation derived from it) does not hold. **2 p**

# Energy minimization & Bayesian estimation [19 points total]

1. Find the solution to the following energy minimization problem **[5 points]**

$$\arg\min_u |Au - f|^2 + \lambda |u - f|^2 \tag{10}$$

where $A \in \mathbf{R}^{n \times n}$ and $u, f \in \mathbf{R}^n$.

**Solution** To find the solution we need to compute the gradient and set it to zero **[2 points]**

$$\nabla E = 2A^\top (Au - f) + 2\lambda(u - f) = 0. \tag{11}$$

By rearranging the equation we obtain **[3 points]**

$$A^\top Au - A^\top f + \lambda(u - f) = 0 \tag{12}$$

$$\left(A^\top A + \lambda I\right) u = (A^\top + \lambda I) f \tag{13}$$

$$u = \left(A^\top A + \lambda I\right)^{-1} (A^\top + \lambda I) f. \tag{14}$$

2. Suppose you are given a collection of images $Y_i \in \mathbb{R}^m, i = 1, \dots n$, and you know $Y_i$ are noisy measurements of an image $X \in \mathbb{R}^m$, such that

$$Y_i = X + \eta, \tag{15}$$

where the noise $\eta \sim \mathcal{N}(0, I)$ is assumed to be of zero mean and unit variance. Derive the maximum likelihood estimate of X.
**Hint 1:** $Y_i \sim \mathcal{N}(X, I)$.
**Hint 2:** The density of the multivariate normal distribution is

$$p(y; \mu, \Sigma) = \det(2\pi\Sigma)^{-\frac{1}{2}} e^{-\frac{1}{2}(y-\mu)^T \Sigma^{-1} (y-\mu)}. \tag{16}$$

**[8 points]**

**Solution** The likelihood of the data given the parameters is

$$p(Y; X) = \prod_{i=1}^{n} (2\pi)^{-\frac{m}{2}} e^{-\frac{1}{2}(Y_i - X)^T (Y_i - X)}. \tag{17}$$

Maximizing this is equivalent to maximizing the log-likelihood:

$$\sum_{i=1}^{n} -\frac{m}{2} \log(2\pi) - \frac{1}{2}(Y_i - X)^T (Y_i - X) \tag{18}$$

We take the derivative w.r.t. $X$ and set it to zero:

$$\sum_{i=1}^{n} X - Y_i = 0. \tag{19}$$

By rearranging the terms, we obtain the estimate

$$X = \frac{1}{n} \sum_{i=1}^{n} Y_i. \tag{20}$$

3. Suppose we are given a task of fitting the parameters of a Gaussian Mixture Model (GMM) $p(x,z)$ to the data $\{x^{(1)},\dots,x^{(m)}\}$ consisting of $m$ independent samples, where $z$ denotes discrete latent variable. Each $z^{(i)}$ identifies the Gaussian from which the sample $x^{(i)}$ was generated.

   (a) Write the data log-likelihood under a Gaussian Mixture Model. [**3 points**]
   **Solution** The data likelihood can be expressed as

$$\begin{aligned}\ell(\theta) &= \sum_{i=1}^{m} \log p(x;\theta) \\ &= \sum_{i=1}^{m} \log \sum_{z} p(x,z;\theta).\end{aligned}$$

   (b) Why do we need the EM algorithm to fit the parameters of GMM? Why do we not simply maximize the likelihood by setting $\nabla_\theta \ell(\theta)$ to 0? [**3 points**]
   **Solution** The main difficulty comes from the fact that variable $z^{(i)}$ are not observed and the posterior does not factorize, making it much harder to compute. [**2 points**] This also complicates the computation of MAP and ML estimates of the parameters. [**1 points**]