LUND UNIVERSITY
FACULTY OF ENGINEERING
MATHEMATICS

TAKE-HOME EXAM
COMPUTER VISION 2019

The solutions should be handed in no later than 48 hours after collecting the exam.
Data for the exam can be downloaded from
`www.maths.lth.se/matematiklth/vision/datorseende/data2019.zip`.
It is not permitted to get help from other persons. Credits can be given for partially solved problems. You may hand in handwritten solutions or pdf-files as well as executable Matlab code. Write your solutions neatly, explain your calculations and specify what Matlab-scripts you have used. Handwritten solutions and printouts should be handed in to the student administration on the 5th floor. Alternatively, you may send your solutions as pdf:s directly to me at `calle@maths.lth.se` *(during office hours when the student administration is open). All your m-files should also be sent directly to me.*

1. *Compute the projections, in regular cartesian coordinates if possible, of the scene points* $(1,1,0)$,$(0,1,2)$ *and* $(1,1,1)$ *in the camera*

$$P = \begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & -1 & 0 & 4 \\ 0 & 0 & -1 & 1 \end{bmatrix}.$$

Also determine the camera center and the principal axis of P.

2. *Consider the cameras* $P_1 = \begin{bmatrix} I & 0 \end{bmatrix}$ *and*

$$P_2 = \begin{bmatrix} 1 & 2 & 0 & 1 \\ 0 & 2 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}.$$

a) *Compute the fundamental matrix* F. *(0.2)*

b) *Suppose* $x = (2,1)$ *is the projection of a point in* P_1. *Which of the points* $(1,0)$,$(2,1)$ *and* $(1,1)$ *could be a projection of the same point in* P_2? *(0.3)*

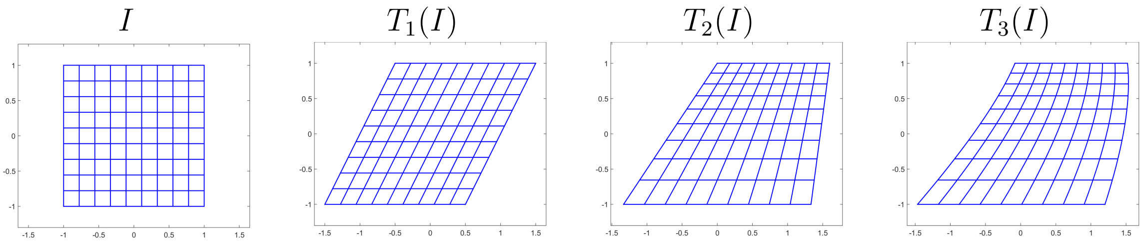
c) *Show that any point on the line*

$$X(\rho) = \begin{pmatrix} 2 \\ 1 \\ 1 \\ \rho \end{pmatrix}.$$

projects to $x = (2,1)$ *in* P_1. *(0.2)*

d) *For each possible correspondence in c) determine the value of* ρ *that gives the correct projections in both* P_1 *and* P_2 *and give a geometric interpretation of* $X(\rho)$ *for this value.* *(0.3)*

3. **a)** *Three transformations $\mathbb{P}^2 \to \mathbb{P}^2$ maps the image shown in I to the images shown in $T_1(I)$, $T_2(I)$ and $T_3(I)$. Determine which of the transformations T_1, T_2 and T_3 that are projective. Are any of them affine-, similarity- or rigid transformations?* *(0.4)*



b) *Show that the three points in $\mathbb{P}^2$ with homogeneous coordinates $(1,0,1)$, $(2,3,1)$ and $(0,-3,1)$ are all lying on one line and determine that line.* *(0.3)*

c) *Let H be the projective transformation $\mathbb{P}^2 \to \mathbb{P}^2$ given by*

$$H = \begin{bmatrix} 1 & 3 & -1 \\ 2 & 2 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

Compute the transformations of the three points in b) and determine if the transformations are on a line. *(0.3)*

4. *Below are two images* `prison11.jpg` *and* `prison2.jpg` *depicting a famous prison. The file* `ex3data.mat` *contains a 3D model X (seen to the right below) and image points $x\{1\}$ and $x\{2\}$ for the two images that have been matched to the 3D model. (NaN in $x\{i\}$ means that the corresponding 3D point was not detected in image i.) The goal of this exercise is to determine the positions and orientations of the cameras that captured the images.*

a) *How many degrees of freedom does the uncalibrated pinhole camera have?*
How many visible 3D points do you need to be able to compute the camera matrix? *(0.2)*

b) *Suppose that the number of mismatched points are roughly 10%. If you use a minimal set of correspondences, how many RANSAC iterations do you need to find an outlier-free set of point correspondences with 99.9% probability?* *(0.3)*

c) *Write a function that computes a camera matrix from a minimal number of correspondences using DLT. Use RANSAC with this function to determine the two camera matrices. A point is considered to be an inlier if its projection is less than 5 pixels from the corresponding image point. Don't forget to make sure that your solutions have visible points in front of them. Plot the 3D model and the camera centers and principal axes in a 3D plot. (0.3)*

d) *Compute the inner parameters of the cameras. Do they seems reasonable?* *(0.2)*

5. *In this exercise the goal is to solve calibrated relative orientation for cameras that are known to lie on a sphere directed towards the origin.*

a) *Show that a camera of the type* $P = \begin{bmatrix} R & z \end{bmatrix}$*, where* R *is a rotation and* $z = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ *has its camera center on the unit sphere and a principal axis that is pointing towards the origin. (0.1)*

b) *Let* $P_1 = \begin{bmatrix} I & z \end{bmatrix}$ *and* $P_2 = \begin{bmatrix} R & z \end{bmatrix}$*. Show that the essential matrix of this camera pair is* $[z - r_3]_\times R$ *where* r_3 *is the third column of* R*. (Hint: First changing coordinates so that* $P_1 = \begin{bmatrix} I & 0 \end{bmatrix}$ *makes it possible to use the standard Essential matrix formula.) (0.2)*

c) *Show that the Essential matrix* $E = [z - r_3]_\times R$ *can be written*

$$E = \begin{bmatrix} e_1 & e_2 & e_3 \\ e_2 & -e_1 & e_4 \\ e_5 & e_6 & 0 \end{bmatrix}.$$

Also show that there are three matrices E_1*,* E_2*,* E_3 *such that the set of matrices of the above form that fulfills the epipolar constraints* $\mathbf{y}^T E \mathbf{x} = 0$ *for three point-matches can be written*

$$E = aE_1 + bE_2 + cE_3.$$

(Hint: We have $[r_3]_\times R = \begin{bmatrix} r_3 \times r_1 & r_3 \times r_2 & r_3 \times r_3 \end{bmatrix}$*, where* r_1*,*r_2 *and* r_3 *are the columns of* R*. Additionally* R *is a rotation.) (0.3)*

d) *The file* `solve_polynomial_system.m` *contains a function that takes the three matrices* E_1, E_2, E_3 *and finds values of* a *and* b *such that* $E = aE_1 + bE_2 + E_3$ *is an Essential matrix. (Note that since the scale of an Essential matrix is arbitrary we can assume that the* c *parameter is* 1*.)*
Use the above function in a RANSAC algorithm (randomly selecting 3 point-matches) to determine an essential matrix for the two images `sphinx1.jpg` *and* `sphinx2.jpg` *shown below. Matches between the two images and camera calibration is available in the variables* x *and* K *in the file* `ex5data.mat`*. An point-match is assumed to be an inlier is the distance between the point and the corresponding epipolar line is less than* 10 *pixels, in both images.*

(0.3)

e) *Extract a camera-matrices from the essential matrix that you obtained in d) and triangulate the 3D points that are inliers. Make sure that your solution has points in front of both cameras. (0.1)*

6. a) *Let* $P_1 = K\begin{bmatrix} I & 0 \end{bmatrix}$ *and* $P_2 = K\begin{bmatrix} R & t \end{bmatrix}$*, where* K *is upper triangular with last row* $\begin{bmatrix} 0 & 0 & 1 \end{bmatrix}$*. Suppose that* $\begin{bmatrix} \pi \\ 1 \end{bmatrix}$*,* $\pi \in \mathbb{R}^3$*, are the homogeneous coordinates of a 3D plane. Furthermore, assume that* $\mathbf{x}_i \sim P_1\mathbf{X}_i$ *and* $\bar{\mathbf{x}}_i \sim P_2\mathbf{X}_i$*, where* $\mathbf{X}_i$*, belongs to the 3D plane and is in front of both cameras for all* $i = 1, ..., n$*. Show that there is a homography* H*, with third row* H_3 *such that*

$$\bar{\mathbf{x}}_i \sim H\mathbf{x}_i, \quad \textit{and} \quad H_3\mathbf{x}_i > 0, i = 1, ..., n.$$

(Hint: Assignment 1 may be of help.) *(0.3)*

b) *The element of the homography* H *in a) depends on the parameters of the plane* π *and we therefore write* $H(\pi)$*. Let*

$$p_i(\pi) = \bar{x}_i - \frac{H_1(\pi)\mathbf{x}_i}{H_3(\pi)\mathbf{x}_i} \quad \textit{and} \quad q_i(\pi) = \bar{y}_i - \frac{H_2(\pi)\mathbf{x}_i}{H_3(\pi)\mathbf{x}_i},$$

where H_1, H_2 *and* H_3 *are the rows of* H *and* $(\bar{x}_i, \bar{y}_i, 1)$ *are homogeneous coordinates of* $\bar{\mathbf{x}}_i$*. Given two fixed numbers* ϵ *and* τ *show that the sets*

$$\mathcal{S}_i = \{\pi \in \mathbb{R}^3; |p_i(\pi)| \leq \tau, H_3(\pi)\mathbf{x}_i \geq \epsilon\}$$

and

$$\mathcal{T}_i = \{\pi \in \mathbb{R}^3; |q_i(\pi)| \leq \tau, H_3(\pi)\mathbf{x}_i \geq \epsilon\}$$

are convex. Is the intersection of all these sets $\mathcal{I} = \cap_{i=1}^{n}(\mathcal{S}_i \cap \mathcal{T}_i)$ *convex? What is the geometric interpretation if these sets? (The number* ϵ *can be assumed to be small so that in practice* $H_3(\pi)\mathbf{x}_i \geq \epsilon$ *is the same as* $H_3(\pi)\mathbf{x}_i \geq 0$*.)* *(0.3)*

c) *The file* `ex6data.mat` *contains cameras and image points for the two datasets shown in the images below. (The cameras* Pa *and the points* xa *are the data for the castle images and* Pa *and* xa *are for the door images.)*

For the castle data set determine if there is a vector $\pi \in \mathcal{I}$ *if* $\epsilon = 10^{-4}$ *and* $\tau = 5$*. If there is compute the corresponding homography* $H(\pi)$ *and transform the points in* $xa\{1\}$ *and plot together with the points in* $xa\{2\}$ *in* `castle2.jpg`*.*

(Hint: The constraint $|p_i(x)| \leq \epsilon$ *is equivalent to* $\begin{cases} p_i(\pi) \leq \epsilon \\ -p_i(\pi) \leq \epsilon \end{cases}$*. Additionally Matlab's* `linprog` *function is useful.)*

For the door data set determine the smallest τ *(integer precision is enough) for which there is a vector* $\pi \in \mathcal{I}$ *if* $\epsilon = 10^{-4}$*. Is the value of* τ *reasonable and why?* *(0.4)*

Good Luck!