Visual Navigation

Winter Term 2014/2015

Prof. Dr. sc. nat. C. Günther

Dr. G. Giorgi

Technische Universität München
Lehrstuhl für Kommunikation und Navigation

Exam

February 18, 2015

- The use of the following material is permitted: lecture slides, referenced books and articles, a pocket (also programmable) calculator. A maximum of four (4) A4 pages of handwritten/typed formulas are allowed. All other items, including books and articles not referenced in the slides, cell phones, laptop etc. . . are prohibited.
- Please indicate your name and registration number on all pages. The problems are described on 5 pages. Please also return those pages.
- The examination is subdivided into 3 areas. Each area typically includes several smaller problems, some of which are dependent.
- The maximum score is reached when a problem is solved. In the case that you cannot solve a problem completely, you have the option to provide a clear and complete description of the solution. A certain credit will also be given for such a description.

I hereby confirm that I have been informed prior to begin of the examination that I have to notify the examination supervisors immediately if sudden illness occurs during the examination. This will be noted in the examination protocol. An application for exam withdrawal has to be filed immediately at the board of examiners being in charge. A medical certificate from one of the physicians acknowledged by the Technische Universität München issued on the same day as the examination must be forwarded without delay. In case the examination is regularly completed despite of illness, a subsequent withdrawal due to illness cannot be accepted. In case the examination is ended due to illness it will not be graded.

Name: ______________________________

Matriculation number: ______________________________

Course of studies: ______________________________

München, ____________ ______________________________
(Date) (Signature)

Problem 1 *Total: 8 points*

Consider a camera with a known calibration matrix:

$$\mathbf{K} = \begin{bmatrix} 400 & 0 & 10 \\ 0 & 500 & 20 \\ 0 & 0 & 1 \end{bmatrix} \tag{1-1}$$

a) (3 points) The projection on the image plane of a point in space $\mathbf{X}$ is obtained as $\mathbf{x} = \mathbf{P}\mathbf{X}$ (using homogeneous coordinates). After rotating the camera around its center $\widetilde{\mathbf{C}}$, while keeping the camera center fixed in space, the projected point on the image shifts from $\mathbf{x}$ to $\mathbf{x}'$. Derive the relation between points $\mathbf{x}$ and $\mathbf{x}'$ in terms of the camera calibration matrix, rotation matrix and camera center.

b) (2 points) Under the same conditions detailed in the previous point, if the rotation matrix is given as

$$\mathbf{R} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix} \tag{1-2}$$

with $\theta = \frac{\pi}{2}$, determine the expected new image coordinates of a feature point which was located at point $(0, 0)^{\mathrm{T}}$ (image coordinates) before applying the rotation.

c) (3 points) A pure translational motion parallel to the image plane is applied to the camera. Furthermore, the camera is only free to rotate about an axis perpendicular to the image plane. Compute the location of the epipoles in image coordinates. You are free to choose the most convenient reference frame.

Problem 2 *Total: 10 points*

A stereo rig is formed by two cameras characterized by exactly the same camera calibration matrix $\mathbf{K}$. The focal distance is 100 pixels, the principal point is in the center of the image plane and the image coordinates are given (in pixels) with respect to the axes x and y as depicted in Figure 2-1). The pixels are squares and no skew distortion is present. The world coordinate frame is centered on the left camera ($\tilde{\mathbf{C}}_l$, with axes oriented as in Figure 2-2). The stereo baseline is 10 centimeters long. The cameras are perfectly aligned (no rectification needed).

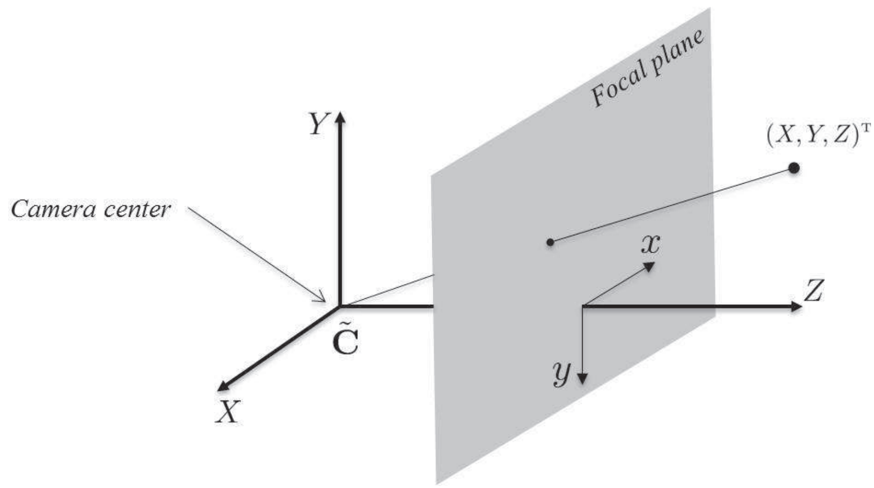

Figure 2-1: Camera local coordinate frame and focal plane for both cameras in the stereo rig.

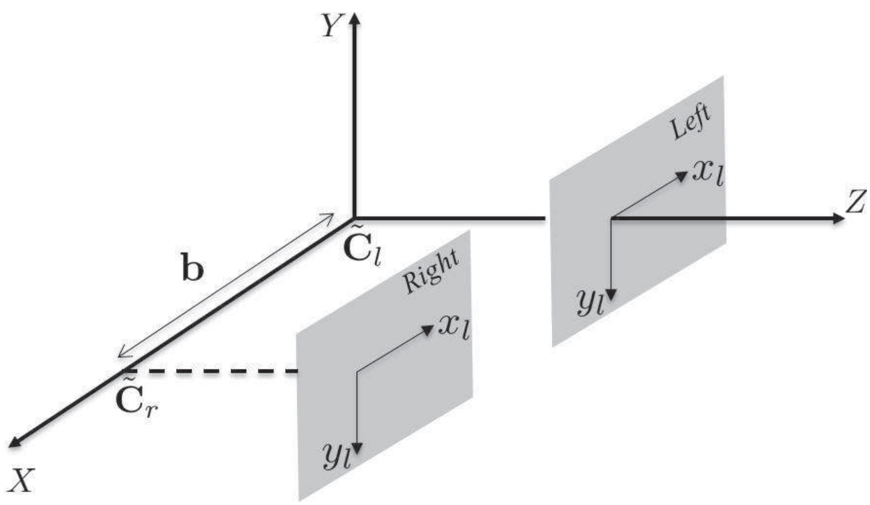

Figure 2-2: The stereo rig.

a) (1 point) Derive the camera calibration matrix as function of the given parameters.

b) (1 point) Derive a coherent construction for both the left and right camera matrices.

c) (2 points) Four distinct points in space are imaged in the left and right cameras. The image homogeneous coordinates of these four points are

$$\begin{aligned}
&\mathbf{x}_l^{(1)} = \begin{pmatrix} 0 \\ 0 \\ 10 \end{pmatrix} \quad \mathbf{x}_l^{(2)} = \begin{pmatrix} -1000 \\ 0 \\ 11 \end{pmatrix} \quad \mathbf{x}_l^{(3)} = \begin{pmatrix} 0 \\ -500 \\ 10 \end{pmatrix} \quad \mathbf{x}_l^{(4)} = \begin{pmatrix} -1000 \\ -500 \\ 11 \end{pmatrix} \quad \begin{bmatrix} \mathrm{px}\cdot\mathrm{m} \\ \mathrm{px}\cdot\mathrm{m} \\ \mathrm{m} \end{bmatrix} \\
&\mathbf{x}_r^{(1)} = \begin{pmatrix} 10 \\ 0 \\ 10 \end{pmatrix} \quad \mathbf{x}_r^{(2)} = \begin{pmatrix} -990 \\ 0 \\ 11 \end{pmatrix} \quad \mathbf{x}_r^{(3)} = \begin{pmatrix} 10 \\ -500 \\ 10 \end{pmatrix} \quad \mathbf{x}_r^{(4)} = \begin{pmatrix} -990 \\ -500 \\ 11 \end{pmatrix} \quad \begin{bmatrix} \mathrm{px}\cdot\mathrm{m} \\ \mathrm{px}\cdot\mathrm{m} \\ \mathrm{m} \end{bmatrix}
\end{aligned} \tag{2-1}$$

with $\mathbf{x}_l^{(i)}$ the homogeneous coordinates of the i-th point in the left image plane and $\mathbf{x}_r^{(i)}$ the homogeneous coordinates of the i-th point in the right image plane.
Reconstruct the four points coordinates in space.

d) (3 points) The image coordinates in both left and right cameras can be measured with an error that can be described as a zero-mean Gaussian variable with standard deviation $\sigma_x = \sigma_y =$0.1 pixels, with σ_x the uncertainty along the image x-axis and σ_y the uncertainty along the image y-axis. There exists a small correlation between the errors in the two axes, measured by a correlation coefficient[1] $\rho = 0.1$.
Derive the covariance matrix $\boldsymbol{\Sigma}_{\mathbf{X}}$ that characterizes the error for one of the points $\mathbf{X}^{(i)}$ that have been reconstructed in the previous point (you are free to choose any of the four points).

e) (3 points) Derive the maximum depth (perpendicular distance from the image plane) of a point in space that enables a reconstruction of the Z-coordinate with an error smaller than $\sigma_Z \leq 10$ centimeters.

[1]The covariance matrix of a vector $\mathbf{z} = (u, v)^{\mathrm{T}}$ is formally built as $\boldsymbol{\Sigma}_{\mathbf{z}} = \begin{bmatrix} \sigma_u^2 & \rho\sigma_u\sigma_v \\ \rho\sigma_u\sigma_v & \sigma_v^2 \end{bmatrix}$

Problem 3 *Total: 12 points*

Consider a self-guiding platform traveling on a monorail at (approximatively) constant speed (see Fig. 3-1). A camera is installed on the platform, pointed downward and looking along the track in front of the platform. The camera tracks landmarks placed at regular intervals on the surface of the rail. The spacing between landmarks is such that at most only one landmark is in the field of view of the camera at any given time. The imaging system provides the position of the landmark along the track *with respect to the current position of the platform*. The internal imaging error is such that the observed landmark position in the platform's own coordinate system can be considered Gaussian, with zero mean and standard deviation equal to σ_m.

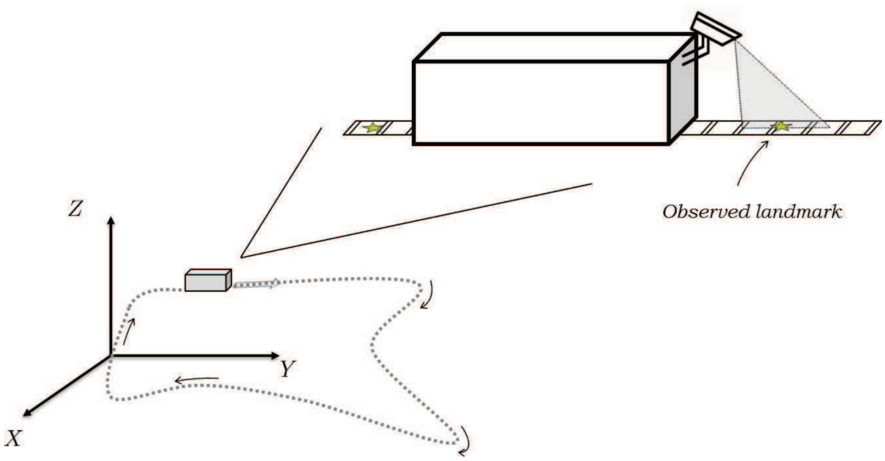

Figure 3-1: A schematic representation of the monorail.

a) (4 points) In order to update the position of the platform in real-time, you are requested to design a Kalman Filter for this specific application. Design (justifying your choices) a dynamic model and an observation model using the *minimum number of variables possible*. Consider the landmarks positions as variables, so the system could be used off-the-shelf without any a-priori exact knowledge of a landmark map, thus being capable of forming its own map with repeated travels along the same monorail. The only information given is that there are M landmarks along the track.

b) (4 points) The probabilistic estimation of the platform and landmarks positions at initialization time $t = 0$ is specified by vector $\boldsymbol{\mu}_0$ and covariance matrix $\boldsymbol{\Sigma}_0$. Assume the initial kinematic state of the platform to be perfectly known, so that

$$\boldsymbol{\Sigma}_0 = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{\mathbf{m}_0} \end{bmatrix} \tag{3-1}$$

with $\boldsymbol{\Sigma}_{\mathbf{m}_0}$ the covariance matrix of an approximate map of the landmarks that is uploaded in the system. This covariance matrix is given as

$$\boldsymbol{\Sigma}_{\mathbf{m}_0} = \frac{\sigma_0^2}{2}\left[\mathbf{I}_M + \mathbf{e}_M\mathbf{e}_M^{\mathrm{T}}\right] = \sigma_0^2 \begin{bmatrix} 1 & \frac{1}{2} & \cdots & \frac{1}{2} \\ \frac{1}{2} & 1 & & \\ \vdots & & \ddots & \\ \frac{1}{2} & & & 1 \end{bmatrix} \tag{3-2}$$

with $\mathbf{e}_M$ a $M \times 1$ vector of ones.
At the following time step (after 1 second), the first landmark is observed. Using the equations derived in the previous problem, derive the functions used to update each of the components of the state vector. Are all the landmarks positions updated? If not, why do they not? If yes, why is it so?

c) (4 points) Following from the previous problem, show that the variance of *all* landmarks decreases after updating, even if only the first landmark was observed. What is causing this effect?