Image- Based Measurement Laboratory
Lab3 - Geometry

Christoph Feichtenhofer

Winter - Term 2017/2018

Teaching Goals

This third laboratory is concerned with the application of cameras to estimate geometric entities. The exercises cover the range from uncalibrated monocular sensors to fully calibrated stereo pairs as frequently used in photogrammetry.

Important Note!

Please strictly adhere to the following rules when working in the laboratory:

- **Never** use a camera without a suitable mount such as a tripod!
- Be careful when re-connecting a camera to the PC! Some interfaces require to completely power down the PC prior to any camera handling.
- **Never** touch the unprotected surface of a sensor!
- Do not modify the aperture, focal length, and focus settings of a calibrated camera.

Exercise 1 - Monocular Calibration

An important prerequisite to the use of cameras in vision-based metrology is the metric calibration of the sensor. Calibration refers to the estimation of the model parameters based on one or more acquisitions of a calibration target. We use a pin-hole model which is comprised of the following parameters:

- Radial and tangential distortion.
- Horizontal and vertical focal lengths.
- Skew angle.
- Principle point.

Using a planar calibration target, perform the following points:

- Acquire $N = 10...15$ images of the calibration target.
- Calibrate the camera intrinsics using the Caltech toolbox.
- Visualise the different geometric constellations.
- Discuss your results.
- Visualise the lens distortion by means of a vector field.
- Undistort one of the calibration images and discuss the results.

Perform again the image stitching task of Lab1 and compare your results.

Exercise 2 - Stereo Calibration

Using the stereo-rig perform the following tasks:

- Calibrate the stereo rig. The intrinsic camera parameters will be provided.
- Search for corresponding points with the help of the epipolar lines.
- Estimate the area of the tracking target.

Exercise 3 - Fundamental matrix estimation

The fundamental matrix $\boldsymbol{F}$ describes the relationship between two views of the same scene. It is a 3×3 rank 2 matrix and can be estimated using a minimum number of 7 point correspondence. In this exercise you will implement the well-known normalized 8 point algorithm ([2], [3]) to estimate the fundamental matrix. To do so, perform the following two steps:

1. Least squares estimation of $\boldsymbol{F}$: for two corresponding points $\boldsymbol{x_i}$ and $\boldsymbol{x'_i}$ the following equation holds: $\boldsymbol{x'_i} \boldsymbol{F} \boldsymbol{x_i} = 0$. Use this fact to set up a linear equation system of the form $\boldsymbol{A}\boldsymbol{f} = \boldsymbol{0}$ and solve for $\boldsymbol{f}$ which should be the vector representation of $\boldsymbol{F}$.
2. Impose the rank 2 constraint of the fundamental matrix: The most convenient way to impose the rank 2 constraint ($\det \boldsymbol{F} = 0$) is to replace the estimate fundamental matrix $\boldsymbol{F}$ with $\hat{\boldsymbol{F}}$. Let $\boldsymbol{F} = \boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^T$ be the singular value decomposition of $\boldsymbol{F}$, with $\boldsymbol{S} = diag(\sigma_1, \sigma_2, \sigma_3)$ being a diagonal matrix with the singular values of $\boldsymbol{F}$. The matrix $\hat{\boldsymbol{F}}$ can be composed according to $\hat{\boldsymbol{F}} = \boldsymbol{U}\hat{\boldsymbol{S}}\boldsymbol{V}^T$, where $\hat{\boldsymbol{S}} = diag(\sigma_1, \sigma_2, 0)$ is the diagonal matrix of the two largest singular values of $\boldsymbol{F}$.

Your task for the exercise is:

- Acquire a stereo image pair of a scene.
- Use 8 corresponding points to estimate $\boldsymbol{F}$.
- Draw epipolar lines to verify the correctness of your estimation.

Exercise 4 - Camera Projection Matrix

The goal of this exercise is to find the *camera projection matrix* $\boldsymbol{P}$ [1] and to decompose it to get the *camera calibration matrix* $\boldsymbol{K}$, the rotation matrix $\boldsymbol{R}$, and the camera center $\tilde{\boldsymbol{C}}$. The projection matrix $\boldsymbol{P}$ is a mapping between homogeneous world coordinates $\boldsymbol{X}_i = (X_i, Y_i, Z_i, 1)$ and homogeneous image coordinates $\boldsymbol{x}_i = (x_i, y_i, w_i)^T$ and thus has the dimension 3×4.

$$\boldsymbol{x}_i = \boldsymbol{P}\boldsymbol{X}_i$$

The matrix $\boldsymbol{P}$ is unique up to scale, thus we use the equation

$$\boldsymbol{x}_i \times \boldsymbol{x_i} = 0 \qquad \Rightarrow \qquad \boldsymbol{x}_i \times \boldsymbol{P}\boldsymbol{X}_i = 0$$

to determine $\boldsymbol{P}$. With $\boldsymbol{P}^{iT}$ being the i-th row of $\boldsymbol{P}$ we can write the cross product as

$$\boldsymbol{x}_i \times \boldsymbol{P}\boldsymbol{X}_i = \begin{pmatrix} y_i \boldsymbol{P}^{3T}\boldsymbol{X}_i - w_i \boldsymbol{P}^{2T}\boldsymbol{X}_i \\ w_i \boldsymbol{P}^{1T}\boldsymbol{X}_i - x_i \boldsymbol{P}^{3T}\boldsymbol{X}_i \\ x_i \boldsymbol{P}^{2T}\boldsymbol{X}_i - y_i \boldsymbol{P}^{1T}\boldsymbol{X}_i \end{pmatrix} \tag{1}$$

And since $\boldsymbol{P}^{iT}\boldsymbol{X}_i = \boldsymbol{X}_i^T \boldsymbol{P}^i$, we get

$$\begin{pmatrix} \boldsymbol{0}^T & -w_i \boldsymbol{X}_i^T & y_i \boldsymbol{X}_i^T \\ w_i \boldsymbol{X}_i^T & \boldsymbol{0}^T & -x_i \boldsymbol{X}_i^T \\ -y_i \boldsymbol{X}_i^T & x_i \boldsymbol{X}_i^T & \boldsymbol{0}^T \end{pmatrix} \begin{pmatrix} \boldsymbol{P}^1 \\ \boldsymbol{P}^2 \\ \boldsymbol{P}^3 \end{pmatrix} = \boldsymbol{0} \tag{2}$$

Because only two rows are linearly independent, we can drop e.g. the third row and get two equations for each point correspondence. And since $\boldsymbol{P}$ has 11 degrees of freedom, we need $5^1/_2$ (6) point correspondences.

To decompose the projection matrix, we need to investigate the structure of $\boldsymbol{P}$. It can be written as

$$\boldsymbol{P} = \boldsymbol{K}\boldsymbol{R}[\boldsymbol{I}| - \tilde{\boldsymbol{C}}]$$

where $\boldsymbol{K}$ is an upper triangular matrix

$$\boldsymbol{K} = \begin{bmatrix} \alpha_x & s & x_0 \\ & \alpha_y & y_0 \\ & & 1 \end{bmatrix}$$

and $\boldsymbol{R}$ is an orthogonal rotation matrix between the camera and the world coordinate frame. On the other hand, $\boldsymbol{P}$ can be written as

$$\boldsymbol{P} = [\boldsymbol{M}| - \boldsymbol{M}\tilde{\boldsymbol{C}}] \tag{3}$$

Using RQ-decomposition we can decompose $\boldsymbol{M} = \boldsymbol{K}\boldsymbol{R}$ into $\boldsymbol{K}$ and $\boldsymbol{R}$.

For this exercise perform the following tasks:

1. Using the 3d target, acquire $n = 6$ images.
2. Compute the projection matrix $\boldsymbol{P}$ using the given MATLAB framework.
3. Decompose the projection matrix $\boldsymbol{P}$ to get:
   - The camera center $\tilde{\boldsymbol{C}}$ in the world coordinate frame.
   - The camera calibration matrix $\boldsymbol{K}$.
   - The rotation matrix $\boldsymbol{R}$ that represents the orientation of the camera coordinate frame.

[1] Multiple View Geometry in Computer Vision. Hartley R. I. and Zisserman A., Cambridge University Press, 2004.

[2] A Computer Algorithm for Reconstructing a Scene from Two Projections. Longuet-Higgins, H. C., 1987

[3] In Defense of the Eight-Point Algorithm. Hartley, Richard I., PAMI, 1997