Computer Vision, Assignment 2
Calibration and DLT

1 Instructions

In this assignment you will study camera calibration, the projective ambiguity, and the DLT method. You will solve the resection and triangulation problems using DLT and compute inner parameters using RQ factorization. In addition you will try out SIFT for feature detection/matching.

Please see Canvas for detailed instructions on what is expected for a passing/higher grade. All exercises not marked **OPTIONAL** are “mandatory” in the sense described on Canvas.

The maximum amount of points for the theoretical exercises in Assignment 2 is 24. 50% of 24 is 12.

2 Calibrated vs. Uncalibrated Reconstruction.

Theoretical Exercise 1 (4 points). Show that when estimating structure and motion (3D points and cameras) simultaneously, under the assumption of uncalibrated cameras (i.e. unknown calibration matrices), there is a projective ambiguity so that the solution can only be recovered up to an additonal unknown projective transformation of 3D space (which cannot be determined using only image projections). That is, if $\mathbf{X}$ are the estimated 3D-points, show that a new solution with identical image projections can always be obtained from $T\mathbf{X}$ for any projective transformation T of 3D space. (Hint: Look at the camera equations.)

For the report: Submit a complete (but short) solution.

Computer Exercise 1. (See Jupyter Notebook)

Theoretical Exercise 2 (6 points). Explain why we can not get the same projective distortions as in Computer Exercise 1 when we use calibrated cameras (i.e. calibration matrices are known). Furthermore, how is the projective ambiguity (described in Exercise 1) affected when the cameras are calibrated – is there still some ambiguity to the reconstruction?

For the report: Submit a short explanation.

3 Camera Calibration

Theoretical Exercise 3 (8 points). Suppose that a camera has got the inner parameters

$$K = \begin{pmatrix} f & 0 & x_0 \\ 0 & f & y_0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{1}$$

Verify that the inverse of K is

$$K^{-1} = \begin{pmatrix} 1/f & 0 & -x_0/f \\ 0 & 1/f & -y_0/f \\ 0 & 0 & 1 \end{pmatrix} \tag{2}$$

and that this matrix can be factorized into

$$K^{-1} = \underbrace{\begin{pmatrix} 1/f & 0 & 0 \\ 0 & 1/f & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{=A} \underbrace{\begin{pmatrix} 1 & 0 & -x_0 \\ 0 & 1 & -y_0 \\ 0 & 0 & 1 \end{pmatrix}}_{=B}. \tag{3}$$

What is the geometric interpretation of the the transformations A and B?

When normalizing the image points of a camera with known inner parameters we apply the transformation K^{-1}. What is the interpretation of this operation? Where does the principal point (x_0, y_0) end up? And where does a point with distance f to the principal point end up?

Suppose that for a camera with resolution 720×400 pixels we have the inner parameters

$$K = \begin{pmatrix} 360 & 0 & 360 \\ 0 & 360 & 200 \\ 0 & 0 & 1 \end{pmatrix}. \tag{4}$$

Normalize the points $(0, 200)$, $(720, 200)$.

What is the angle between the viewing rays projecting to these points?

Show that the camera $K[R\ t]$ and the corresponding normalized version $[R\ t]$ have the same camera center and principal axis (viewing direction). HINT: For a camera matrix P, note that the null space of P defines the camera center, and let the last row P_{31}, P_{32}, P_{33} define the principal axis. NOTE: There is a special family of camera matrices known as affine cameras, for which the last row is 0, and for which the principal axis can not be determined in this way. However, you can disregard this, as the calibration matrix structure given above can not yield any affine camera.

For the report: Complete solution.

4 RQ Factorization and Computation of K

Theoretical Exercise 4 (4 points). Consider an upper triangular matrix

$$K = \begin{pmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{pmatrix}, \tag{5}$$

where the diagonal elements of K are positive. Verify by matrix multiplication that

$$KR = \begin{pmatrix} aR_1^T + bR_2^T + cR_3^T \\ dR_2^T + eR_3^T \\ fR_3^T \end{pmatrix}, \text{ where } R = \begin{pmatrix} R_1^T \\ R_2^T \\ R_3^T \end{pmatrix}. \tag{6}$$

If

$$P = \begin{pmatrix} 2400\sqrt{2} & 0 & 800\sqrt{2} & 4000 \\ 700\sqrt{2} & 2800 & -700\sqrt{2} & 4900 \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} & 3 \end{pmatrix}, \tag{7}$$

what is R_3 (R is an orthogonal matrix) and f?

Using the second row of the camera matrix, and the fact that $R_3 \perp R_2$, $||R_3|| = ||R_2|| = 1$ compute R_2,d,e. (Hint: If $v = dR_2 + eR_3$ how do you compute the coefficient e?)

Similarly, using the first row compute, R_1,a,b and c. What is the focal length, skew, aspect ratio, and principal point of the camera?

For the report: Complete solution.

5 Direct Linear Transformation DLT

Theoretical Exercise 5 (**OPTIONAL, 15 optional points**). Show that the linear least squares system

$$\min_v ||Mv||^2 \tag{8}$$

always has the minimum value 0. In the DLT algorithm we use the least squares system

$$\min_{||v||^2=1} ||Mv||^2 \tag{9}$$

to remove the zero solution. Show that if M has the singular value decomposition $M = U\Sigma V^T$ then

$$||Mv||^2 = ||\Sigma V^T v||^2 \tag{10}$$

and

$$||V^T v|| = 1 \text{ if } ||v||^2 = 1. \tag{11}$$

Consider a new optimization problem

$$\min_{||\tilde{v}||^2=1} ||\Sigma\tilde{v}||^2. \tag{12}$$

Explain why (12) gives the same minimal value as (9). How can you obtain a solution to the first problem from the second (determine v^* given $\tilde{v}^*$)? Also, explain why there are always at least two solutions to these (equivalent) problems.

Finally, for the $m \times n$ matrix M, prove that the last column of V is an explicit solution to (9), in the case when $\text{rank}(M) < n$. Consider what happens when $m \geq n$, as well as when $m < n$. HINT: The singular values are in decreasing order on the diagonal of Σ. NOTE: The assumption $\text{rank}(M) < n$ is in fact not necessary but the general case is more difficult to prove, see lecture notes.

For the report: Complete solution.

Theoretical Exercise 6 (2 points). When using DLT it is often advisable to normalize the points before doing computations. Note that normalizing the points in this context should not be confused with the calibration of cameras and image points, which can also be referred to as normalization. Instead, this is just a technique which in practice yields higher quality solutions when using the DLT method.

Suppose the image points $\mathbf{x}$ are normalized by the mapping N by

$$\tilde{\mathbf{x}} \sim N\mathbf{x} \tag{13}$$

and that we compute a camera matrix in the new coordinate system, that is, we obtain a camera $\tilde{P}$ that solves

$$\tilde{\mathbf{x}} \sim \tilde{P}\mathbf{X}. \tag{14}$$

How do you compute the camera P that solves the original problem

$$\mathbf{x} \sim P\mathbf{X} \tag{15}$$

from $\tilde{P}$?

For the report: It is a very simple exercise, just give the formula.

Computer Exercise 2. (See Jupyter Notebook)

6 Feature Extraction and Matching using SIFT

Computer Exercise 3. (See Jupyter Notebook)

7 Triangulation using DLT

Computer Exercise 4. (See Jupyter Notebook)