Computer Vision Exercise 1: Camera Calibration

Soomin Lee (leesoo@student.ethz.ch)

October 2, 2020

1. Overview

In this exercise, two camera calibration algorithms are introduced: the Direct Linear Transform (DLT) algorithm and the Gold Standard algorithm. The first task is to normalize the data, which is then used in both algorithms. Data normalization is performed in a way that ensures the input points to have zero mean and unit mean distance to the origin. As a result of this step, transformation matrices T and U are obtained in the following format:

$$T = \begin{bmatrix} \sigma_{2d} & 0 & \bar{x} \\ 0 & \sigma_{2d} & \bar{y} \\ 0 & 0 & 1 \end{bmatrix}, \quad U = \begin{bmatrix} \sigma_{3d} & 0 & 0 & \bar{x} \\ 0 & \sigma_{3d} & 0 & \bar{y} \\ 0 & 0 & \sigma_{3d} & \bar{z} \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{1}$$

The DLT algorithm uses the fact that $x_i \times PX_i = 0$ for each i-th point. It builds up a matrix using $n \geq 6$ points by stacking up A_i in Equation (2), and then try to find a solution of the linear system. Therefore, the algorithm minimizes the algebraic error. The projection matrix P_n for the normalized points can be de-normalized before factorizing it into K, R, and t using $P = T^{-1}P_nU$.

$$A_i = \begin{bmatrix} 0_{1\times4} & -X_i^T & y_iX_i^T \\ X_i^T & 0_{1\times4} & -x_iX_i^T \end{bmatrix} \tag{2}$$

The Gold Standard algorithm follows a similar procedure as DLT algorithm, but it minimizes the geometric error E in Equation (3) instead of the algebraic error. Consequently, it has an additional step where it optimizes the projection matrix parameters, using the reprojection error as the cost function. It uses the result from the DLT algorithm for the initial value.

$$E = \frac{1}{N}\sum_{i=1}^{N} d(\hat{x}_i, \hat{P}\hat{X}_i)^2 \tag{3}$$

2. Result - Intrinsic Parameters / Average Reprojection Errors

The values of the intrinsic parameters in Equation 4 are obtained from the DLT algorithm, and the values of the intrinsic parameters in Equation 5 are obtained from the Gold Standard algorithm.

Note that input coordinates (X, Y, Z) were specified in mm.

$$\begin{aligned} P &= \begin{pmatrix} 2.6297 & -0.1222 & -0.4232 & -528.7235 \\ 0.8907 & 1.4436 & 1.8440 & -690.3393 \\ 0.0009 & 0.0014 & -0.0006 & -0.6104 \end{pmatrix}, \\ K &= \begin{pmatrix} 1339.5 & 6.8 & 800.6 \\ 0 & 1346.3 & 581.3 \\ 0 & 0 & 1 \end{pmatrix}, \quad R = \begin{pmatrix} -0.8468 & 0.5319 & -0.0074 \\ -0.1695 & -0.2830 & -0.9440 \\ -0.5042 & -0.7981 & 0.3298 \end{pmatrix}, \quad t = \begin{pmatrix} 16.7217 \\ 145.8485 \\ 357.2872 \end{pmatrix} \end{aligned} \tag{4}$$

$$P = \begin{pmatrix} 2.6222 & -0.1252 & -0.4209 & -527.6452 \\ 0.8897 & 1.4434 & 1.8396 & -689.2170 \\ 0.0009 & 0.0014 & -0.0006 & -0.6094 \end{pmatrix},$$
$$K = \begin{pmatrix} 1338.4 & 7.8 & 796.0 \\ 0 & 1345.3 & 580.7 \\ 0 & 0 & 1 \end{pmatrix}, \quad R = \begin{pmatrix} -0.8472 & 0.5312 & -0.0071 \\ -0.1701 & -0.2838 & -0.9437 \\ -0.5033 & -0.7983 & 0.3308 \end{pmatrix}, \quad t = \begin{pmatrix} 17.7827 \\ 145.9882 \\ 356.9209 \end{pmatrix} \tag{5}$$

The average reprojection errors are calculated as in Equation 3, and the values are 8.7927 and 7.8773 in pixels respectively. As expected, the error of the Gold Standard algorithm is lower than that of the simple DLT algorithm.

3. Visualization of the Result

In Figure 1 and Figure 2, the hand-clicked points are marked with red circles while the reprojected 3D points are marked with green crosses, obtained using the DLT algorithm and the Gold Standard algorithm respectively. As the reprojection error from the DLT algorithm is already small, the difference between the two figures are almost negligible.

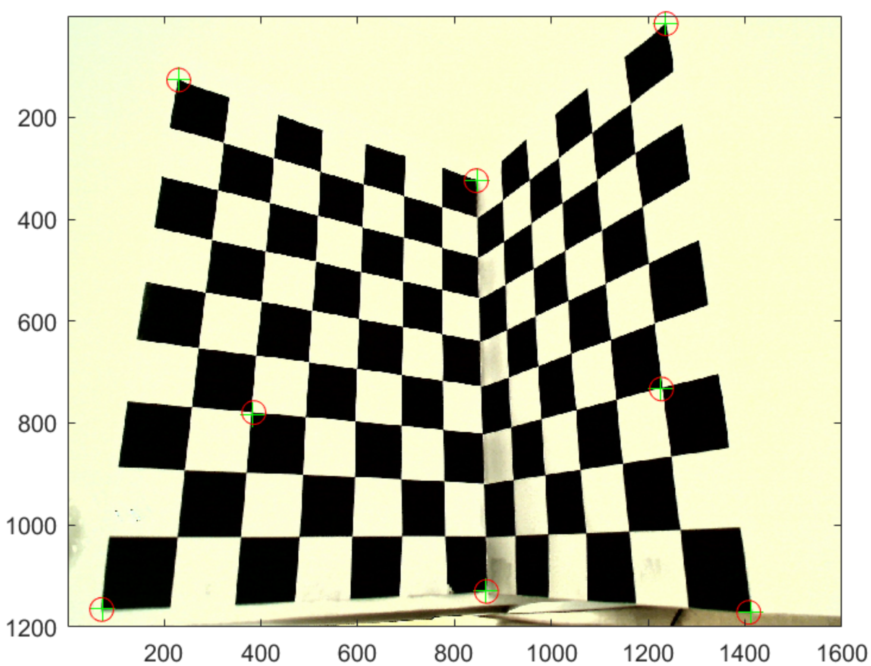

Figure 1: Visualization of the hand-clicked points (red circle) and the reprojected 3D points (green cross) obtained using the DLT algorithm

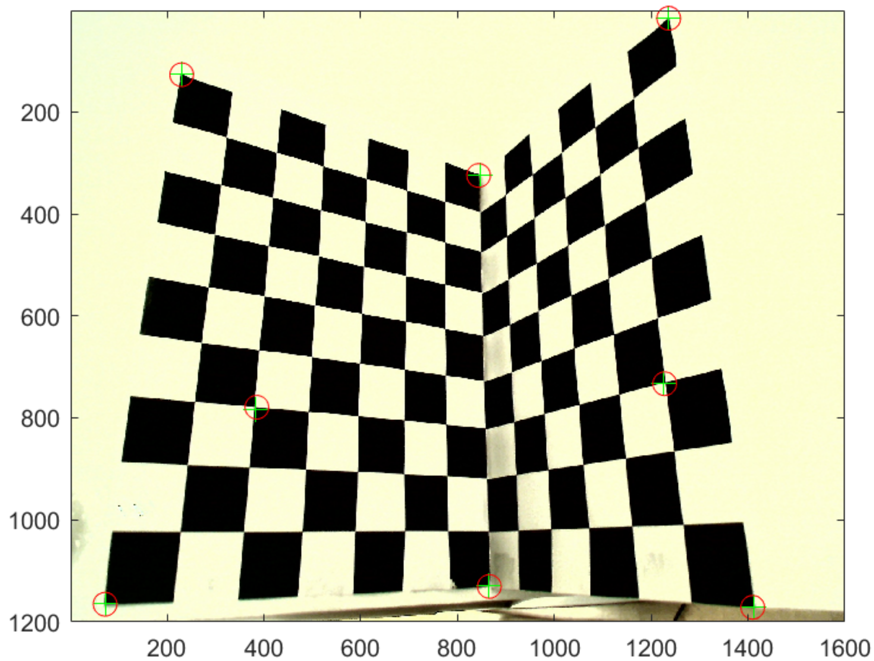

Figure 2: Visualization of the hand-clicked points (red circle) and the reprojected 3D points obtained (green cross) using the Gold Standard algorithm