Homework 2

Programming Assignment

In the code, I choose 300 out of 713 points in the central region. First, construct matrix $\mathbf{A}$ to calculate matrix $\mathbf{H}$:

$$\begin{bmatrix} x_w^{(1)} & y_w^{(1)} & 1 & 0 & 0 & 0 & -u_1 x_w^{(1)} & -u_1 y_w^{(1)} & -u_1 \\ 0 & 0 & 0 & x_w^{(1)} & y_w^{(1)} & 1 & -v_1 x_w^{(1)} & -v_1 y_w^{(1)} & -v_1 \\ x_w^{(2)} & y_w^{(2)} & 1 & 0 & 0 & 0 & -u_2 x_w^{(2)} & -u_2 y_w^{(2)} & -u_2 \\ 0 & 0 & 0 & x_w^{(2)} & y_w^{(2)} & 1 & -v_2 x_w^{(2)} & -v_2 y_w^{(2)} & -v_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_w^{(n)} & y_w^{(n)} & 1 & 0 & 0 & 0 & -u_n x_w^{(n)} & -u_n y_w^{(n)} & -u_n \\ 0 & 0 & 0 & x_w^{(n)} & y_w^{(n)} & 1 & -v_n x_w^{(n)} & -v_n y_w^{(n)} & -v_n \end{bmatrix} \begin{bmatrix} h_{11} \\ h_{12} \\ h_{13} \\ h_{21} \\ h_{22} \\ h_{23} \\ h_{31} \\ h_{32} \\ h_{33} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Since $\mathbf{B} = \mathbf{K}^{-T} K^{-1}$ and:

$$\begin{cases} \mathbf{h_1^T B h_1} - \mathbf{h_2^T B h_2} = 0 \\ \mathbf{h_1^T B h_2} = 0 \end{cases}$$

We can compute matrix $\mathbf{B}$ using the same method. Then, $\mathbf{K}$ can be solved by cholesky decomposition.
The reprojection error of i-th image can be expressed as:

$$\frac{1}{n} \sum_{i=1}^{n} ||\mathbf{u_i} - \mathbf{H x_i}||^2$$

The result is shown below. According to the result, the f_X, f_y, s, o_x, o_y of the intrinsic matrix is $751.4, 696.0, 2.4, 360.7, 276.4$ respectively. And the 4 reprojection errors are 1.5, 4.9, 5.8, 21.7 respectively(4 images detected only). The implement details are shown in the code.

Written Assignment

(a) The sum of the squard errors can be written as:

$$E(\mathbf{A}, \mathbf{T}) = \sum_{i=1}^{N} (\mathbf{Y_i} - \mathbf{A X_i} - \mathbf{T})^T (\mathbf{Y_i} - \mathbf{A X_i} - \mathbf{T})$$

```
PS E:\Homework> & E:/download/Anaconda/envs/pytorch_dl/python.exe e:/Homework/CV_HW2/hw.py
Camera Calibration Matrix:
[[751.4019      2.3793697 360.6617   ]
 [  0.        696.00464   276.4355   ]
 [  0.          0.          1.       ]]
Camera Calibration Matrix by OpenCV:
Camera Matrix:
 [[730.46122482   0.         456.97841613]
 [  0.         681.51960987 341.21083089]
 [  0.           0.           1.        ]]
Reprojection error: [1.4538009, 4.893909, 5.798035, 21.728924]
```

The partial derivatives of $\mathbf{A}, \mathbf{T}$ equals to zero:

$$\begin{cases} \dfrac{\partial E}{\partial \mathbf{A}} = \displaystyle\sum_{i=1}^{N} 2(\mathbf{A}\mathbf{X_i} + \mathbf{T} - \mathbf{Y_i})\mathbf{X_i}^T = \mathbf{0} \\ \dfrac{\partial E}{\partial \mathbf{T}} = \displaystyle\sum_{i=1}^{N} 2(\mathbf{A}\mathbf{X_i} + \mathbf{T} - \mathbf{Y_i}) = \mathbf{0} \end{cases}$$

We can learn from the equation that:

$$\mathbf{T} = \frac{1}{N}\sum_{i=1}^{N}(\mathbf{Y_i} - \mathbf{A}\mathbf{X_i}) = \hat{\mathbf{Y}} - \mathbf{A}\hat{\mathbf{X}}$$

Therefore, we have:

$$\sum_{i=1}^{N}(\mathbf{A}\mathbf{X_i} - \mathbf{A}\hat{\mathbf{X}} + \hat{\mathbf{Y}} - \mathbf{Y_i})\mathbf{X_i^T} = \mathbf{0}$$

$$\mathbf{A}(\sum_{i=1}^{N}(\mathbf{X_i} - \hat{\mathbf{X}})\mathbf{X_i^T}) = \sum_{i=1}^{N}(\mathbf{Y_i} - \hat{\mathbf{Y}})\mathbf{X_i^T}$$

The above equation can be written as:

$$\mathbf{A}\mathbf{X}\mathbf{X^T} = \mathbf{Y}\mathbf{X^T}$$

The final transformation is therefore given by:

$$\begin{cases} \mathbf{A}^* = (\mathbf{Y}\mathbf{X^T})(\mathbf{X}\mathbf{X^T})^{-\mathbf{1}} \\ \mathbf{T}^* = \hat{\mathbf{Y}} - \mathbf{A}^*\hat{\mathbf{X}} \end{cases}$$

(b) Since $\mathbf{A} \in \mathbb{R}^{3\times 3}$ and $\mathbf{T} \in \mathbb{R}^3$, we have 12 unknowns. And for each correspondence, we have $\mathbf{Y_i} = \mathbf{A}\mathbf{X_i} + \mathbf{T}$, which contains 3 constraints. Therefore, 4 is the minimum number of correspondences needed to estimate the transformation.