AI3604 – Computer Vision
Fall 2023
Homework #2
Due: 11:59 PM, Thursday, November 30, 2023

Programming Assignment (85 points)

In this course assignment, you will be required to perform camera calibration. Given a series of chessboard images, you will need to deduce the camera's focal length, principal point. Only basic Python, Numpy, and OpenCV functions are allowed, unless otherwise specified. This rule applies in particular with the use of OpenCV. Basic image IO functions such as *cv2.imread*, *cv2.imwrite* etc. and corner detect functions (*e.g. cv2.findChessboardCorners*) are allowed. If you are unsure of an allowable function, please ask on OC or ask TA before assuming! You will get no credit for using a "magic" function to answer any questions where you should be answering them. When in doubt, ASK! For all functions that take or return images, make sure to handle the actual arrays. Do not pass filenames around.

In terms of grading, we would not like to judge students by how good the programming results are. Our purpose is to help the students to learn more from the homework, so do not be afraid of your programming results. **As long as you're serious about it, we will give you a fair grade.**

The dataset is available in: https://jbox.sjtu.edu.cn/l/z14zf9

The standard calibration process can be summarized as follows:

1. Define real world coordinates of 3D points using checkerboard pattern of known size.
2. Capture the images of the checkerboard from different viewpoints.
3. Find chessboard corners as well as finding the pixel coordinates (u, v) for each 3D point in different images
4. Find camera parameters, the 3D points, and the pixel coordinates using linear algebra.

The final results should include:

1. The camera's intrinsic parameters, including the focal length, principal point (distortion parameters are not required.)

2. The projection error corresponding to each image (which is the error after projecting the 3D points back onto the pixel plane; the smaller the error, the more accurate the estimation of the intrinsic parameters).

Notice:

1. For simplification, you can pass the points in XY plane as (0,0), (1,0), (2,0), ... which denotes the location of points. In this case, the results will be in the scale of size of chess board square.
2. To minimize the impact of the distortion parameters, please try to calculate the intrinsic parameters using the central region of the images as much as possible.

Written Assignment (15 points)

1. Let $\{\mathbf{X}_i \in \mathbb{R}^3\}_{i=1}^N$ be a set of point in $\mathbb{R}^3$ that are transformed by a 3D transformation (**A**,**T**), where $\mathbf{A} \in \mathbb{R}^{3\times3}$ is a nonsingular matrix and $\mathbf{T} \in \mathbb{R}^3$ is a translational vector, to produce another set of points $\{\mathbf{Y}_i \in \mathbb{R}^3\}_{i=1}^N$. Suppose that the transformed points $\mathbf{Y}_i$ are corrupted by noise $\mathbf{E}_i$, i.e., $\mathbf{Y}_i = \mathbf{A}\mathbf{X}_i + \mathbf{T} + \mathbf{E}_i$ for all $i = 1, \ldots\ldots, N$.
   a. Show that the transformation (**A**,**T**) that minimizes the sum of the squared errors
   $$E(\mathbf{A}, \mathbf{T}) = \sum_{i=1}^{N} \left|\left|\mathbf{Y}_i - \mathbf{A}\mathbf{X}_i - \mathbf{T}\right|\right|_2^2$$
   is given by $\mathbf{T}^* = \widehat{\mathbf{Y}} - \mathbf{A}^*\widehat{\mathbf{X}}$, $\mathbf{A}^* = (\mathbf{Y}\mathbf{X}^\intercal)(\mathbf{X}\mathbf{X}^\intercal)^{-1}$,where $\widehat{\mathbf{X}} = \sum \mathbf{X}_i / N$, $\mathbf{X} = \left[\mathbf{X}_1 - \widehat{\mathbf{X}}, \ldots, \mathbf{X}_N - \widehat{\mathbf{X}}\right]$ and similarly for $\widehat{\mathbf{Y}}$ and **Y**.

   b. Show that 4 is the minimum number of correspondences needed to estimate the transformation.