Guillaume Drui
Jakub Zíka
Tymofii Voitekh

Computer Vision

NYCU

Homework 1 Report

Introduction

Camera calibration is a fundamental process in computer vision and photogrammetry that involves determining the internal characteristics of a camera (intrinsic parameters) and its position and orientation in the world (extrinsic parameters). These parameters are essential for accurately mapping 3D points in the real world to 2D points in an image plane, which is crucial for applications such as 3D reconstruction, augmented reality, robotic navigation, and motion tracking.

In this project, we implement a camera calibration algorithm from scratch, without relying on high-level functions from computer vision libraries like OpenCV. By using images of a chessboard pattern captured from different angles, we can extract known 3D world coordinates and their corresponding 2D image projections. The chessboard provides a convenient and precise way to obtain a grid of points that are evenly spaced and easily detectable in images.

The motivation behind manually implementing the camera calibration process is to gain a deeper understanding of the mathematical foundations and computational techniques involved. By delving into the details of homography estimation, intrinsic and extrinsic parameter computation, and matrix decomposition, we reinforce our comprehension of key concepts in computer vision.

Implementation procedure

<u>step 1: homography matrix</u>

We are given points in the world coordinates (U,V) and points in the image plane (u,v)
Since all points lie in a plane, their W component is 0 in world coordinates we can thus delete 3rd column of extrinsic matrix like it is done in the lecture.

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \sim \begin{bmatrix} f/s_x & 0 & o_x \\ 0 & f/s_y & o_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & \cancel{r_{13}} & t_1 \\ r_{21} & r_{22} & \cancel{r_{23}} & t_2 \\ r_{31} & r_{32} & \cancel{r_{33}} & t_3 \end{bmatrix} \begin{bmatrix} U \\ V \\ W \\ 1 \end{bmatrix}$$

Starting from this we know that:

$$\begin{bmatrix} u' \\ v' \\ w' \end{bmatrix} = H \begin{bmatrix} U \\ V \\ 1 \end{bmatrix}$$

Where the unknown is H

$$H = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix}$$

Writing this out we get that

$$
\begin{aligned}
u' &= h_{11}U + h_{12}V + h_{13} \\
v' &= h_{21}U + h_{22}V + h_{23} \\
w' &= h_{31}U + h_{32}V + h_{33}
\end{aligned}
$$

To obtain the pixel coordinates (u,v) from the homogeneous coordinates (u′,v′,w′), we need to convert them to inhomogeneous coordinates by dividing by the third coordinate w′

$$
\begin{aligned}
u &= \frac{u'}{w'} = \frac{h_{11}U + h_{12}V + h_{13}}{h_{31}U + h_{32}V + h_{33}} \\
v &= \frac{v'}{w'} = \frac{h_{21}U + h_{22}V + h_{23}}{h_{31}U + h_{32}V + h_{33}}
\end{aligned}
$$

Which is equal to

$$
\begin{aligned}
u(h_{31}U + h_{32}V + h_{33}) &= h_{11}U + h_{12}V + h_{13} \\
v(h_{31}U + h_{32}V + h_{33}) &= h_{21}U + h_{22}V + h_{23}
\end{aligned}
$$

Rewriting these in a standard linear form, they become:

$$
\begin{aligned}
h_{11}U + h_{12}V + h_{13} - h_{31}uU - h_{32}uV - h_{33}u = 0 \\
h_{21}U + h_{22}V + h_{23} - h_{31}vU - h_{32}vV - h_{33}v = 0
\end{aligned}
$$

In our code these equations are directly stored into the matrix A which can be solved to retrieve H

<u>step 2: Matrix B</u>

using the following definitions from the lecture:

$$
\begin{aligned}
\mathbf{h}_1^\top \mathbf{K}^{-\top}\mathbf{K}^{-1}\mathbf{h}_2 &= 0 \\
\mathbf{h}_1^\top \mathbf{K}^{-\top}\mathbf{K}^{-1}\mathbf{h}_1 &= \mathbf{h}_2^\top \mathbf{K}^{-\top}\mathbf{K}^{-1}\mathbf{h}_2
\end{aligned}
$$

$\mathbf{B} := \mathbf{K}^{-\top}\mathbf{K}^{-1}$ **is symmetric and positive definite**

we first solving the first equation we get:

$$
h_1^T \cdot (Bh_2) = \begin{bmatrix} h_{11} & h_{12} & h_{13} \end{bmatrix} \begin{bmatrix} B_{11}h_{21} + B_{12}h_{22} + B_{13}h_{23} \\ B_{12}h_{21} + B_{22}h_{22} + B_{23}h_{23} \\ B_{13}h_{21} + B_{23}h_{22} + B_{33}h_{23} \end{bmatrix}
$$

Which written out equals to:

$$
h_1^T B h_2 = h_{11}B_{11}h_{21} + h_{11}B_{12}h_{22} + h_{11}B_{13}h_{23} + h_{12}B_{12}h_{21} + h_{12}B_{22}h_{22} + h_{12}B_{23}h_{23} + h_{13}B_{13}h_{21} + h_{13}B_{23}h_{22} + h_{13}B_{33}h_{23}
$$

Solving the second equation follows analogously. We then put those equations into a matrix V to find the unknown parameters of B where for every homography we append the rows:

$$
V_{12} = \begin{bmatrix} h_{11}h_{21} & h_{11}h_{22} + h_{12}h_{21} & h_{12}h_{22} & h_{13}h_{21} + h_{11}h_{23} & h_{13}h_{22} + h_{12}h_{23} & h_{13}h_{23} \end{bmatrix}
$$

$$
V_{11} - V_{22} = \begin{bmatrix} h_{11}^2 - h_{21}^2 & 2(h_{11}h_{12} - h_{21}h_{22}) & h_{12}^2 - h_{22}^2 & 2(h_{11}h_{13} - h_{21}h_{23}) & 2(h_{12}h_{13} - h_{22}h_{23}) & h_{13}^2 - h_{23}^2 \end{bmatrix}
$$

we then use SVD to find the best values for B step 3: Extracting the intrinsic and extrinsic matrix

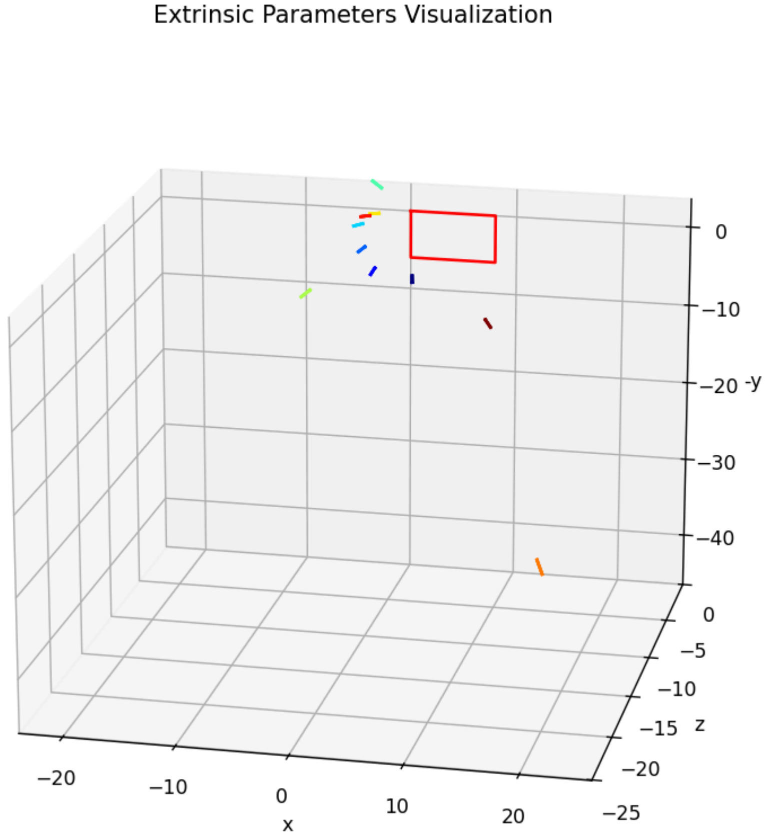
From B we use Cholesky decomposition to extract the intrinsic matrix K. Furthermore we use the following equations to find the extrinsic parameters

$$\mathbf{r}_1 = \lambda \mathbf{K}^{-1}\mathbf{h}_1$$
$$\mathbf{r}_2 = \lambda \mathbf{K}^{-1}\mathbf{h}_2$$
$$\mathbf{r}_3 = \mathbf{r}_1 \times \mathbf{r}_2$$
$$\mathbf{t} = \lambda \mathbf{K}^{-1}\mathbf{h}_3$$
$$\lambda = 1/||\mathbf{K}^{-1}\mathbf{h}_1||$$

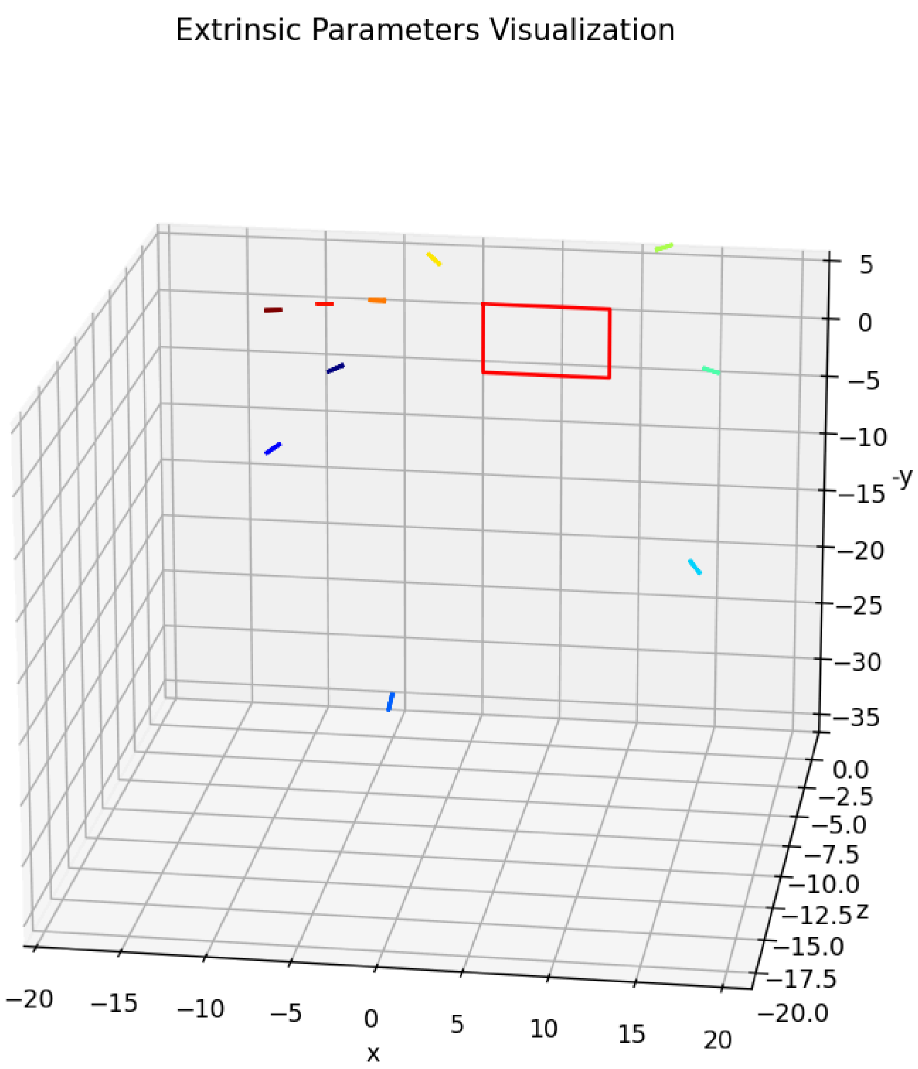
in our code this is done using simple np operations.

Experimental result

Using the provided data here is the 3D evaluation of the extrinsic parameters



For our own images we get the following 3D visualization:



Discussion

The calibration results demonstrate a good level of accuracy in estimating both the intrinsic and extrinsic parameters. However, several factors influenced the precision of the calibration:

- Image Quality: The resolution and clarity of the chessboard images significantly affected the detection of feature points.
- Pattern Detection: Slight variations in the detection of corners due to lighting conditions or perspective distortions introduced minor errors in the homography estimation.

Conclusion

The manual implementation of the camera calibration process deepened our understanding of the underlying mathematical concepts. We successfully determined the camera's intrinsic and extrinsic parameters without relying on high-level library functions. This experience highlighted the importance of precise data acquisition and careful computational practices.

Work assignment plan between team members

The workload has been divided as follows:

- Guillaume Drui:
  - Code review
  - Report writing
  - documenting the mathematical derivations

- Jakub Zíka:
  - Code review
  - Testing the code with our own chessboard pictures
  - Analyzing the calibration results

- Tymofii Voitekh:
  - code writing
  - Implementation of the calibration algorithm