CV HW 1: AutoCalib

Jesdin Raphael Computer Science
Worcester Polytechnic Institute
Email: jraphael@wpi.edu

***Abstract*—In this assignment, I learned about camera calibration. Here I estimate the parameters of a camera such as the focal length, distortion coefficients, and principle points. These parameters were estimated based on the paper "A Flexible New Technique for Camera Calibration" by Zhengyou Zhang.**

I. Introduction

Multiple images of a printed chessboard with a square side length of 21.5 is taken with focus locked. These images are used to calculate camera calibration matrix K which is given as follows

$$\mathbf{K} = \begin{bmatrix} f_x & 0 & \gamma \\ 0 & f_y & c_x \\ 0 & 0 & 1 \end{bmatrix} \tag{1}$$

and the radial distortion parameters $\mathbf{k} = [k1, k2]$

In this report, I estimate the camera calibration parameters $fx, fy, cx, cy, k1, k2$ by implementing the paper "A Flexible New Technique for Camera Calibration" by Zhengyou Zhang [1].

II. Estimation of Camera Intrinsic Matrix

We want to estimate the Camera Intrinsic Matrix which is given by

$$\mathbf{A} = \begin{bmatrix} \alpha & \gamma & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2}$$

A. Homography Estimation

To estimate the Homography for each image we first need the world co-ordinates and the chessboard corner co-ordinates. The chessboard corner points can be found using the $cv2.findChessboardCorners$ method [2]. Whereas, the world coordinates can be easily calculated by using the $SquareLenght$ and the number of rows and columns in the chess board. The corners detected are shown in Fig 1.

Using the detected corners the Homography between the world coordinates and the image coordinates is estimated.

B. Closed Form Solution

Let,

$$\mathbf{B} = A^{-T}A^{-1} \equiv \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{12} & B_{22} & B_{23} \\ B_{13} & B_{23} & B_{33} \end{bmatrix} \tag{3}$$

B is symmetric and can be defined by a 6D vector:

$$b = [B_{11}, B_{12}, B_{22}, B_{13}, B_{23}, B_{33}]^T \tag{4}$$

Let the ith column vector of H be $h_i = [h_{i1}, h_{i2}, h_{i3}]^T$. Then, we have

$$h_i^T B h_j = v_{ij}^T b \quad (7)$$

with

$$\begin{aligned} v_{ij} = [&h_{i1}h_{j1}, \\ &h_{i1}h_{j2} + h_{i2}h_{j1}, \\ &h_{i2}h_{j2}, \\ &h_{i3}h_{j1} + h_{i1}h_{j3}, \\ &h_{i3}h_{j2} + h_{i2}h_{j3}, \\ &h_{i3}h_{j3}]^T. \end{aligned} \tag{5}$$

Which can be rewritten as 2 homogeneous equations in b. Let

$$\begin{bmatrix} \mathbf{v}_{12}^T \\ (\mathbf{v}_{11} - \mathbf{v}_{22})^T \end{bmatrix} b = 0. \tag{6}$$

If n images of the model plane are observed, by stacking n such equations as (8) we have

$$\mathbf{V}b = 0,$$

where $\mathbf{V}$ is the matrix formed by stacking all the vectors $\mathbf{v}_{ij}$.

We then solve for b using SVD.

C. Extraction of the Intrinsic Parameters from Matrix B

Matrix B, shown in Eqn (3), is estimated up to a scale factor, i.e., $B = \lambda A^{-T} A$ with λ an arbitrary scale. we can uniquely extract the intrinsic parameters from matrix B using the equation in Eqn (7).

$$\begin{aligned} v_0 &= \frac{B_{12}B_{13} - B_{11}B_{23}}{B_{11}B_{22} - B_{12}^2} \\ \lambda &= B_{33} - \frac{[B_{13}^2 + v_0(B_{12}B_{13} - B_{11}B_{23})]}{B_{11}} \\ \alpha &= \sqrt{\frac{\lambda}{B_{11}}} \\ \beta &= \frac{\lambda B_{11}}{B_{11}B_{22} - B_{12}^2} \\ \gamma &= -\frac{B_{12}\alpha^2\beta}{\lambda} \\ u_0 &= \frac{\gamma v_0}{\beta} - \frac{B_{13}\alpha^2}{\lambda} \end{aligned} \tag{7}$$

D. Extraction of Extrinsic parameters from Intrinsic Parameters

Now that we have estimated the Intrinsic Parameters A, the extrinsic parameters for each image can be easily computed by using Eqn (2) and Eqn (8)

$$\begin{aligned}
\mathbf{r}_1 &= \lambda A^{-1}\mathbf{h}_1 \\
\mathbf{r}_2 &= \lambda A^{-1}\mathbf{h}_2 \\
\mathbf{r}_3 &= \mathbf{r}_1 \times \mathbf{r}_2 \\
\mathbf{t} &= \lambda A^{-1}\mathbf{h}_3 \\
\text{with} \quad \lambda &= \frac{1}{\|A^{-1}\mathbf{h}_1\|} = \frac{1}{\|A^{-1}\mathbf{h}_2\|}.
\end{aligned} \tag{8}$$

E. Optimization of Parameters

Since we assumed that the camera has minimal distortion we can assume that $k = [0, 0]^T$ for a good initial estimate.

We have the initial estimates of K, R, t, k, now we want to minimize the geometric error defined as given below:

$$\sum_{i=1}^{N}\sum_{j=1}^{M} ||x_{i,j} - \hat{x}_{i,j}(K, R_i, t_i, X_j, k)||$$

Formally, the optimization problem is as follows [3]:

$$\arg\min_{f_x, f_y, c_x, c_y, k_1, k_2} \sum_{i=1}^{N}\sum_{j=1}^{M} ||x_{i,j} - \hat{x}_{i,j}(K, R_i, t_i, X_j, k)|| \tag{9}$$

Using $scipy.optimize$ we minimize the loss function in Eqn (9) to get optimal values.

III. Results

The calibration matrix before and after optimization is shown below:

$$A_{initial} = \begin{bmatrix} 2061.23357 & -0.776635813 & 767.30258 \\ 0.0 & 2042.400242 & 1342.184745 \\ 0.0 & 0.0 & 1.0 \end{bmatrix}$$

$$A_{optimal} = \begin{bmatrix} 2061.23357 & -0.776635813 & 767.30256 \\ 0.0 & 2042.400242 & 1342.18477 \\ 0.0 & 0.0 & 1.0 \end{bmatrix}$$

The distortion coefficients before and after optimization are shown below:

$$K_{initial} = \begin{bmatrix} 0.0 & 0.0 \end{bmatrix}$$
$$K_{optimal} = \begin{bmatrix} 0.0162150 & -0.0909954 \end{bmatrix}$$

The mean projection error after optimization is 0.66122

The projected points are shown in Fig 2.

References

[1] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 11, pp. 1330–1334, 2000.

[2] O. Contributors. (2022) Camera calibration with opencv. Accessed on: February 6, 2024. [Online]. Available: https://docs.opencv.org/3.4/dc/dbb/tutorial_py_calibration.html

[3] R. . R. Perception and M. Learning, "Rbe 549 spring 2024 - homework 1," https://rbe549.github.io/spring2024/hw/hw1/#report, accessed: Feb 2024.

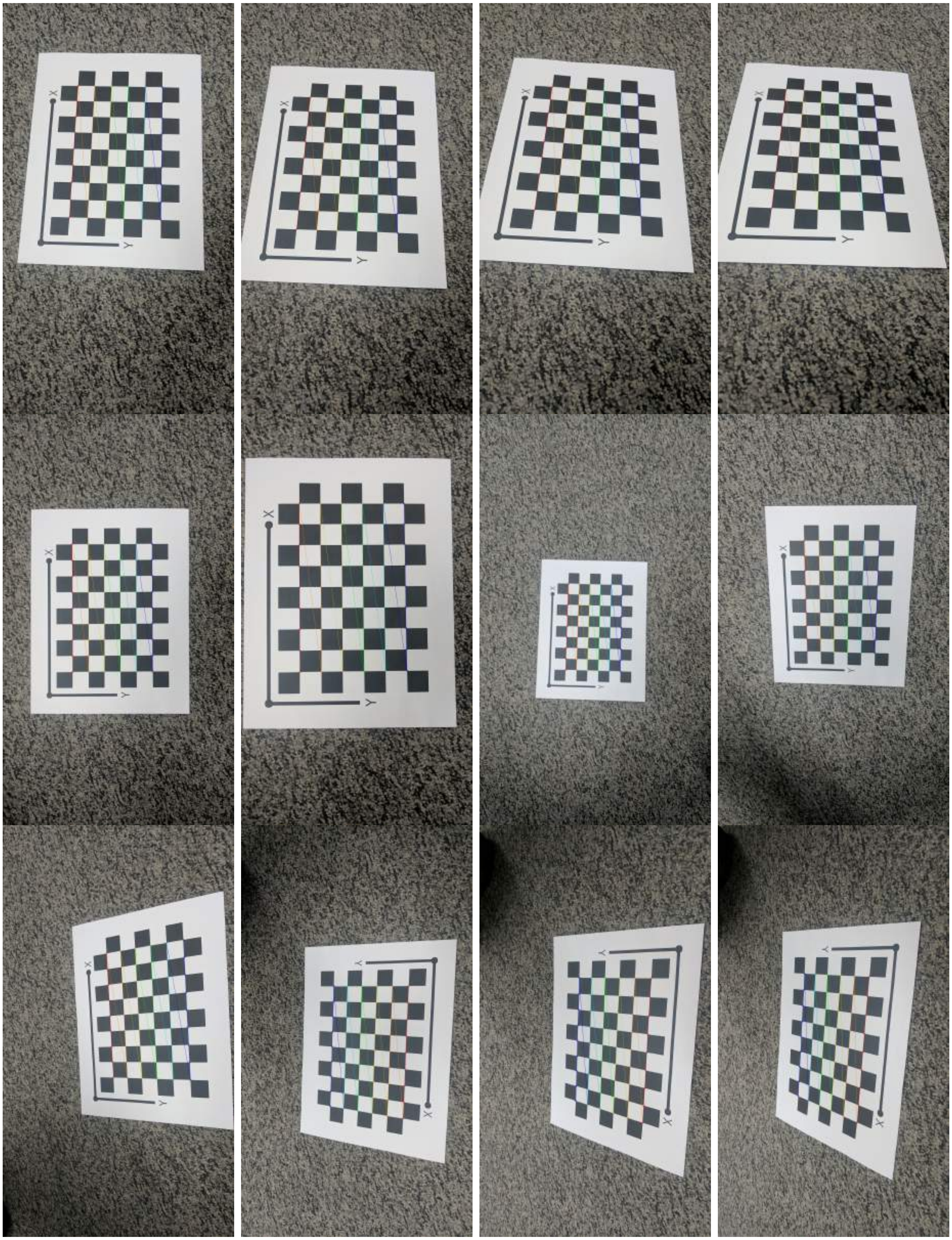

Fig. 1: Detected corner for images

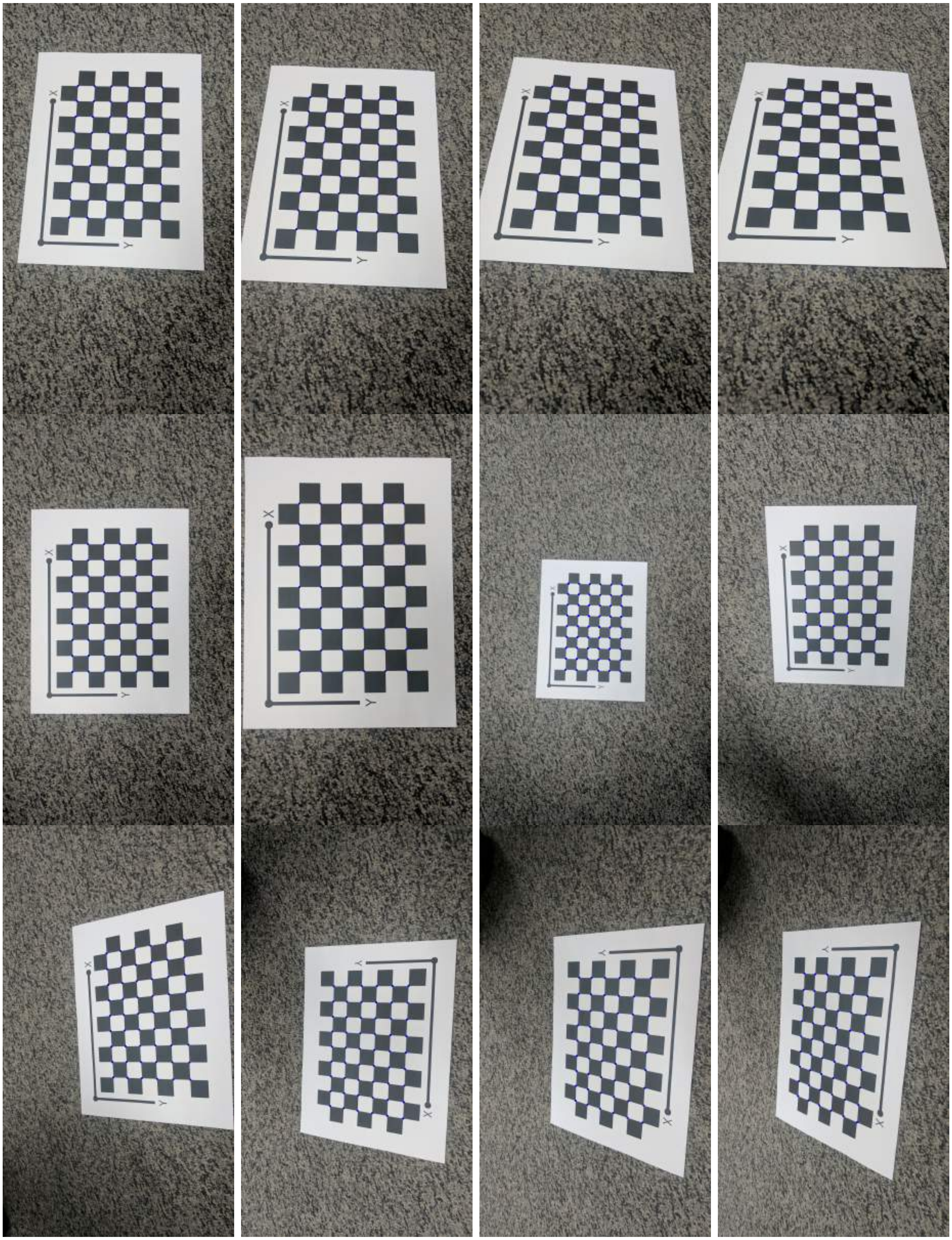

Fig. 2: Reprojected corners