Camera Calibration and Real-World 2D Measurement Using Perspective Projection

Akshat Namdeo
CSC 8830 - Computer Vision
Homework Assignment
Week 2

I. Introduction

This work evaluates camera calibration and real-world two-dimensional measurement using a single calibrated image. The implementation is carried out in Python using OpenCV and follows the pinhole camera model and perspective projection framework presented in the course videos. The objective is to calibrate a smartphone camera using a planar checkerboard target and then use the calibrated camera parameters to estimate the real-world dimensions of a planar object from a single image captured at a known distance.

II. Camera Calibration

Camera calibration is performed using a planar checkerboard pattern displayed on a laptop screen. The checkerboard consists of 13 by 10 squares, resulting in 12 by 9 detectable inner corner points, consistent with OpenCV's checkerboard detection convention.

Five images of the checkerboard are captured using a smartphone camera from different viewing angles and distances to ensure sufficient geometric diversity. All checkerboard points lie on a single plane, and their world coordinates are defined as

$$\mathbf{X}_w = [x_w, y_w, 0, 1]^T$$

The mapping from world coordinates to pixel coordinates is given by

$$\mathbf{u} = \mathbf{K}[\mathbf{R} \mid \mathbf{t}]\mathbf{X}_w$$

where $\mathbf{K}$ is the intrinsic camera matrix, $\mathbf{R}$ and $\mathbf{t}$ describe the camera pose, and $\mathbf{u} = [u, v, 1]^T$ denotes homogeneous pixel coordinates.

The intrinsic matrix has the form

$$\mathbf{K} = \begin{bmatrix} f_x & 0 & o_x \\ 0 & f_y & o_y \\ 0 & 0 & 1 \end{bmatrix}$$

where f_x and f_y are the effective focal lengths in pixel units and (o_x, o_y) is the principal point.

Calibration quality is evaluated using the mean reprojection error, defined as the root-mean-square pixel distance between detected image points and projected world points.

III. Real-World 2D Measurement

After calibration, a single image of a planar object is used to estimate real-world dimensions. The test object is a book placed upright against a wall so that its surface is approximately parallel to the image plane. The image is captured using the same calibrated camera from a measured distance of approximately 2.2 meters.

The image is first undistorted using the estimated intrinsic parameters and distortion coefficients. Four corner points of the book are manually selected in pixel coordinates, as suggested in the videos, to isolate the geometric measurement task.

Under perspective projection, image coordinates satisfy

$$u = \frac{f_x X}{Z}, \quad v = \frac{f_y Y}{Z}$$

For two points lying on the same planar surface at constant depth, real-world displacements are related to pixel displacements by

$$\Delta X = \frac{Z}{f_x}\Delta u, \quad \Delta Y = \frac{Z}{f_y}\Delta v$$

Pixel distances between the selected corner points are computed along the width and height of the book and converted into real-world dimensions using the calibrated focal lengths and known camera-to-object distance.

IV. Experimental Validation

The estimated dimensions are validated using a reference image that shows the true dimensions of the book. The true book size is 7.6 inches in width and 10 inches in height, corresponding to 0.193 meters by 0.254 meters.

Using the calibrated camera parameters and a camera-to-object distance of 2.2 meters, the estimated dimensions are approximately 0.174 meters in width and 0.224 meters in height. This corresponds to percentage errors of approximately 10 percent for width and 12 percent for height.

The remaining discrepancy is attributed to uncertainty in the camera-to-object distance, manual pixel coordinate selection, slight deviation from perfect fronto-parallel alignment, and residual lens distortion. The results are consistent with the expected behavior of perspective projection under the assumptions outlined in the videos.