Robotic Vision Project

Eline Marie Håve
Peder Tinholt Kristensen

May 5, 2023

1 Introduction

In this small project we attempt depth estimation of traffic cones using stereo camera. The goals are:

- Intrinsic calibration of each camera
- Extrinsic calibration of the stereo camera setup
- Detection of apriltags placed on traffic cones
- Depth estimation by triangulating
- Attempt feature matching using SIFT

2 Work

2.1 Intrinsic Camera Calibration

In the project, we used two iPhones to create a stereo camera. We first performed intrinsic calibration of each camera. A sample image from this process is shown in Figure 1.

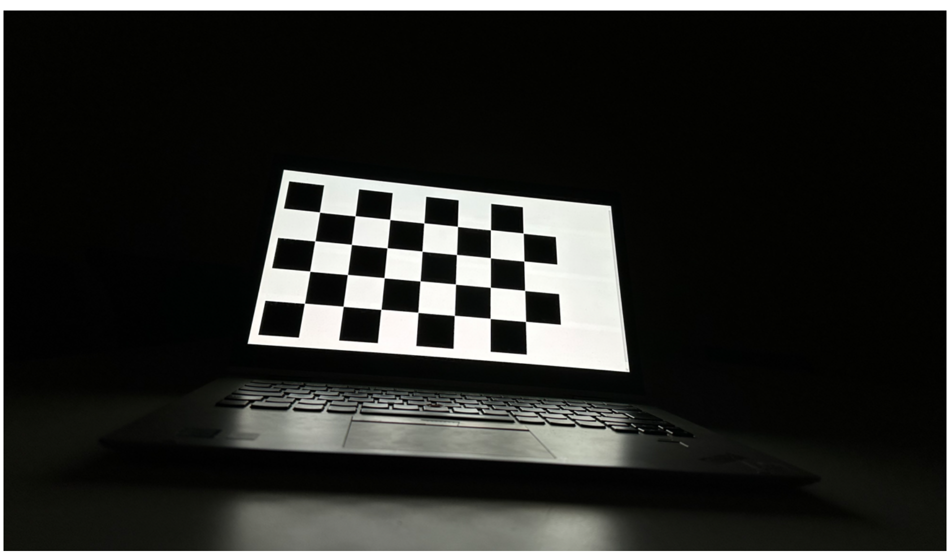

Figure 1: Sample image from intrinsic calibration

After the images were taken, we used a modified version of the intrinsic calibration script given in homework 5.

We then took calibration images using a checkerboard as shown in Figure 3.

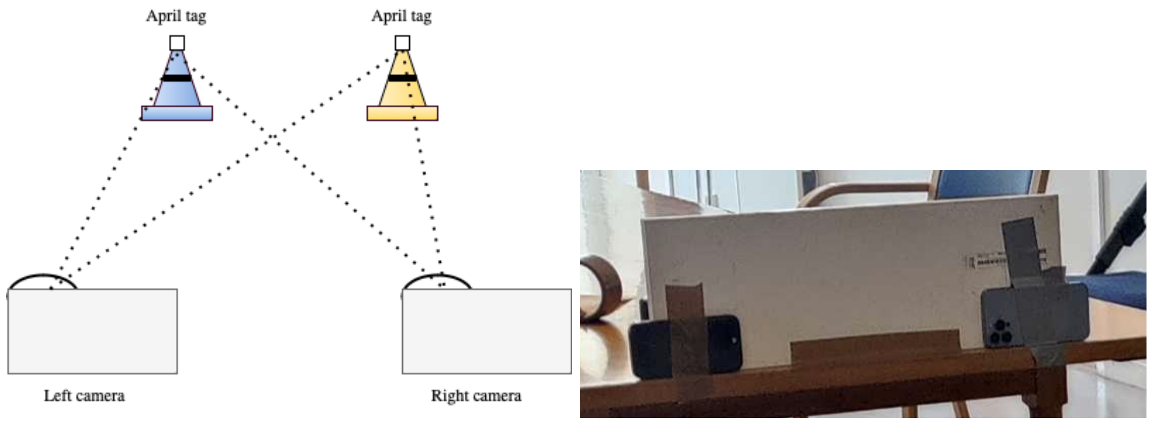

Figure 2: Stereo camera setup

Figure 3: Example of images used

2.2 April Tag Detection

The ID, tag-family and location of the april tags were detected by using the `readAprilTag` function in MatLab. Markers were plotted in the four corners of the tag, as can be seen in Figure 5. The location of the markers in in the image were exported as a csv file, to be used for triangulating the position of the cones.

2.3 Triangulation

During our initial attempt at triangulation the marker points exported from from the april tag detection using the `cv2.triangulatePoints()` function, we observed a significant discrepancy between the triangulated points and the ground truth. Upon inspection of the extrinsic calibration data, it became apparent that the cameras had moved between captures, and that a larger number of images should have been used for the extrinsic calibration.

In subsequent attempts, we captured more images with the cameras remaining stationary. However, it was difficult to identify the calibration board in quite a few images, and as a result, several image pairs had to be excluded from the calibration process. Because of this, its natural that we observe discrepancy between the triangulated points and the ground truth, as depicted in Figure 6. This may be attributed to the limited number of calibration images that we were ultimately able to use. However, studying Figure 6 we do see that the triangulated points somewhat correspond to their ground truths.

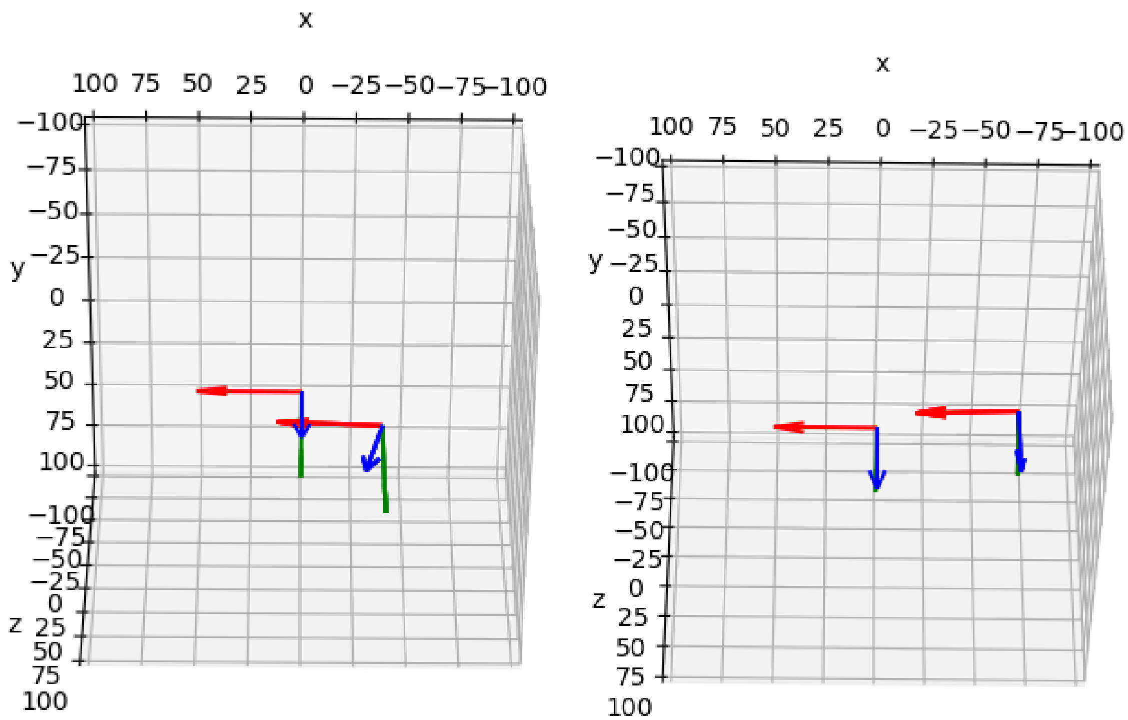

Figure 4: Estimated calibration setups. Right image shows the final setup

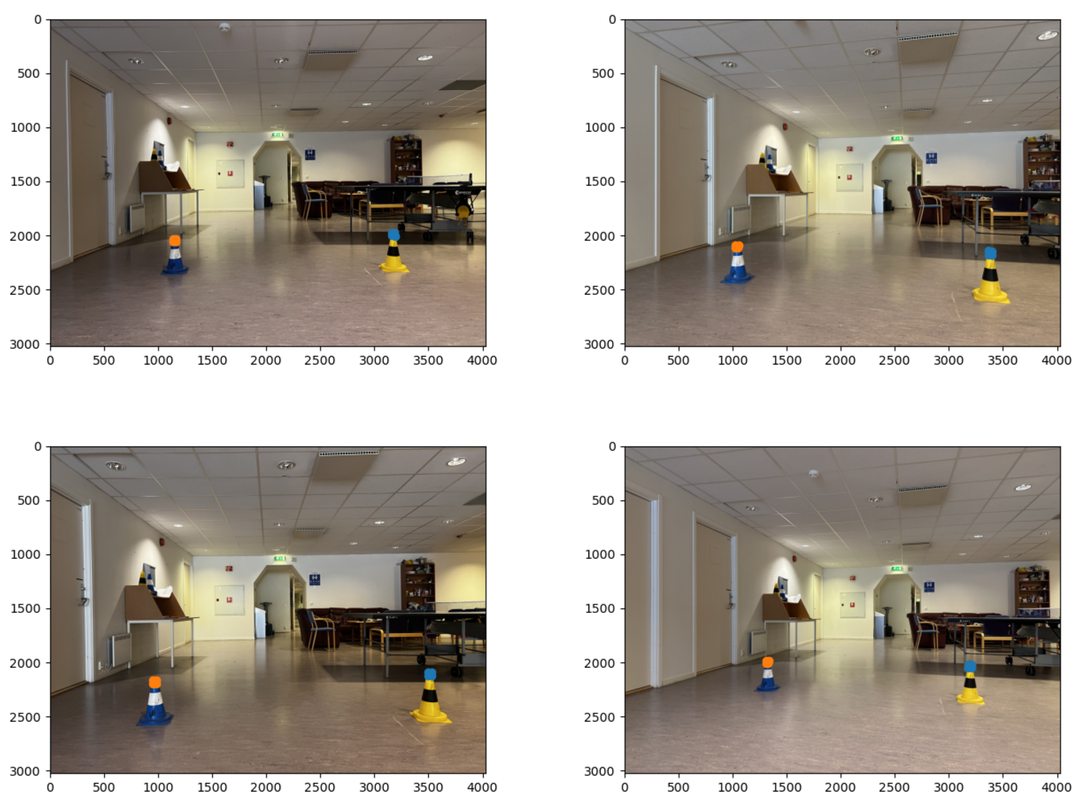

Figure 5: Detected april tags

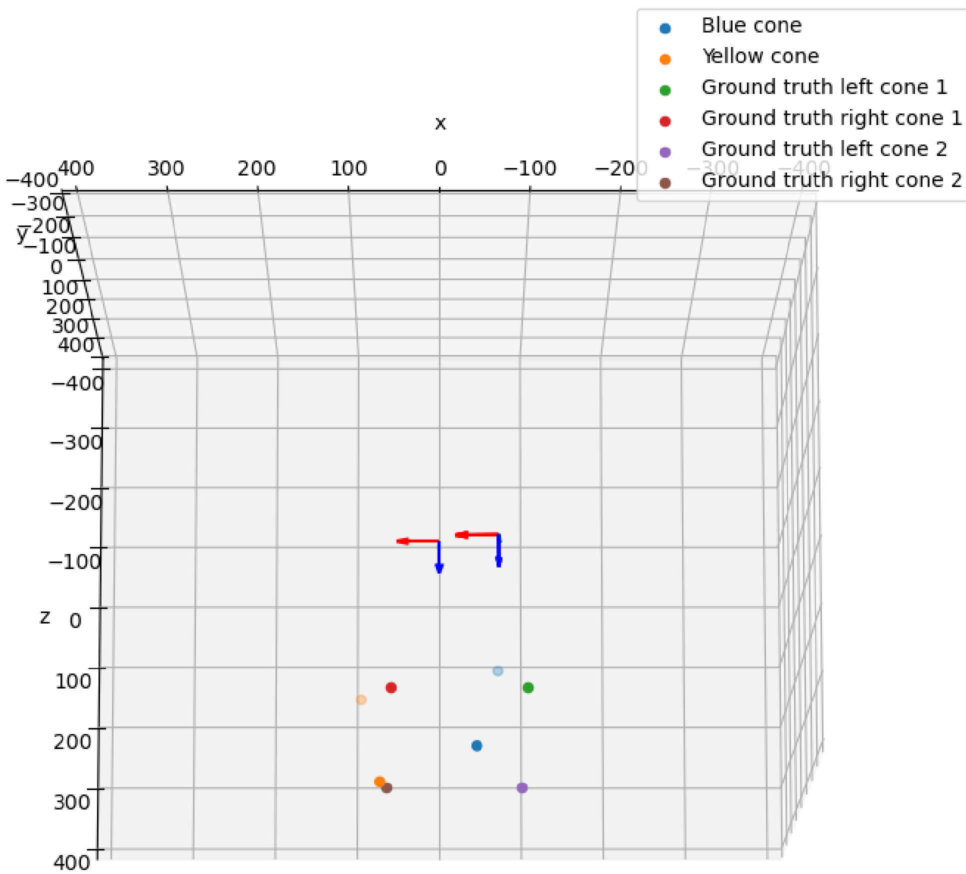

Figure 6: Estimated camera poses, triangulated points and ground truths

2.4 Feature matching

We also attempted to use SIFT, `cv2.BFMatcher` feature matcher in OpenCV, and Lowe's ratio test for feature matching across two images. The results are depicted in Figure 7.

As we can see from the figure, the feature matching works reasonably well although there are some bad matches.

Figure 7: Feature matching using SIFT

Appendix

Figure 8: We tried with our limited time(,: