Computer Vision brief

Jacob Warrer

September 16, 2025

This document will serve as a short brief of the practical applications of Computer Vision technology and related topics performed by Jacob Warrer during his employment at Phase One and enrollment at SDU. The use of Computer Vision, hereafter CV, is assumed to be through the OpenCV [2].

1 Calibration Laboratory

During my employment at Phase One, one of my primary tasks was the development of a semi-automated camera calibration lab [4] for measuring the intrinsic and extrinsic camera parameters. This lab would consist of a wall of Circular Coded Targets(CCT) [1] to which a camera is pointing. The camera is mounted in a three axis gimbal in order to automate the capture of images of the targets in multiple poses, see Figure 1.

In order to deploy this in production a full-stack application was developed to handle capturing of images, moving the camera according to the current location in the lab and post-processing of images. A more detailed description of the vision challenges will be described in the coming subsections.

1.1 Pose estimation

In order for the camera to determine the position and pose in the laboratory it was decided, through theoretical discussion of cost/benefit, that an image based pose estimation would be the best solution for this application. The image based pose estimation used a single from the camera when this was positioned in the lab,

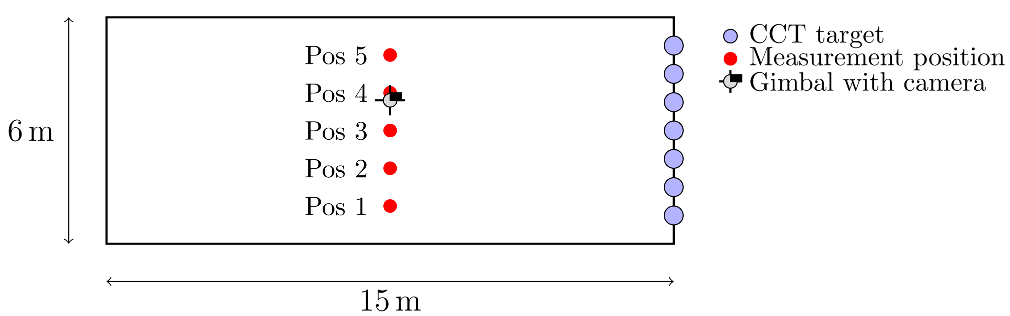

Figure 1: Top-down schematic of the work area with CCT targets on the wall, measurement positions, and a camera/gimbal on a central column.

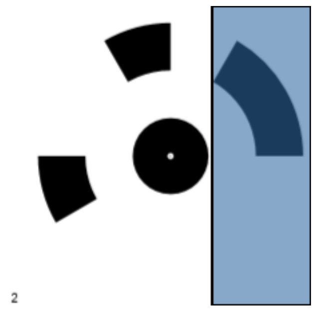

Figure 2: Obfuscated CCT target

pointing roughly at the CCT targets. After the image capture, the 2D positions of the targets were derived from the image data alongside the active focal length. If a multi-camera system was used, an image from each camera would be processed, but only the camera with the best CCT detection would be used.

From the array of 2D point, the focal length and a model of the targets in a world coordinate frame, the pinhole camera model could be used to deduct the position and orientation of the camera. Using the CV function solvePnP [3], the object frame (CCT target) in relation to the camera frame could be determined. From this and the use of transformation chains a rough estimation of the camera pose could be determined. With the pose of the camera and the gimbal and the CCT targets defined in world coordinates, the necessary angle/axis rotation could be calculated and fed to the gimbal in order for it to point the cameras optical axis to the center of the target structure.

1.2 Spatial target filtering

Another challenge that was presented during the development, was to develop a precise and accurate method of filtering the targets. This was necessary due to structural overlap of the targets, where some of the target codes would be obfuscated. This is illustrated by the blue rectangle in Figure 2 where the center is visible but the right side of the binary code circle is hidden. The obfuscation would lead to the detection engine to read the binary target code wrong.

Instead of having to reprint all targets and spend time to refit the laboratory, an investigation into spatial filtering of targets were commenced. In the initial investigative phase, it was noticed that the correct targets estimated point, in a given frame would be inside the center dot of the CCT target. With this assumption in hand the radius of the center of the CCTs had to be determined.

In order to find the radius of a given target with position $\mathbf{p} = \begin{bmatrix} x \; y \end{bmatrix}$ the size of the necessary local Region Of Interest(ROI) was determined empirically from existing images. From $\mathbf{p}$, a vertical and horizontal slice of pixels in the positive direction was sampled. From the sampled pixels, the derivative was calculated, and the edge was to be determined. The final radius was the floored average pixel count from $\mathbf{p}$ to the edge between the vertical and horizontal direction. This was wrapped in a function that would sample $\mathbf{n}$ number of targets in the given image and return the average radius for **n** targets.

With the radius in hand the detection was simplified to a square around **p** with length $2r$. If a target detected in a given image had another estimated target within the detection square, the target was wrongfully detected and thereby discarded.

As this algorithm was to be used in a production environment it was important for it to not only be correct, but also efficient. This resulted in an optimization effort reducing the runtime significantly by analysis of the input data and computational cost of operations.

2 Hardware testing

The internship project was founded in the Phase One MSI system [5]. This system utilize a wide spectrum camera with low-pass and band-pass filters and different light sources of various wavelengths. This project revolved around researching new methods of changing a filter. During this project a system weight limitation was investigated and in order to determine focus accuracy. To find the weight limits while maintaining the accuracy of the focus system, CV was used in conjunction with a checkerboard pattern. The general idea of the experiment, was to move the focus system between the maximum and a set-point and evaluate the sharpness and drift of the system.

3 Focus estimation

During the investigative phase of writing the bachelor project, CV was used to determine the point of maximum sharpness. This was done as a bench experiment where a camera would be mounted in a static environment pointed at multiple Siemens stars. A script was developed to capture and analyze images captured throughout the entire focus range. During the development of the script multiple contrast functions was tested in order to find the sweet spot between accuracy and noise. The goal was to use the optical system to determine the exact focus distance in order to evaluate the proposed lidar distance estimation method.

4 University degree

During the studies at SDU, several semester projects involved programming and control of diverse robotic systems" Course work extends to both traditional Computer Vision through a dedicated course on the OpenCV framework but also interpretation of other sensor modalities, such as various TOF sensors. Coursework during the studies at SDU also included a specialized course on deep neural networks, with an emphasis on image classification and segmentation.

4.1 Pick and throw

The most notable project involving computer vision was a pick-and-throw robot cell, where a camera mounted above the robot cell needed to be linked to a 6-axis collaborative robot.

This challenge required expertise in several computer vision disciplines, such as camera calibration, object detection and localization, and static and dynamic analysis of the data. Furthermore the 2D-location of the object, had to be translated into the frame of the robot to facilitate the required inverse kinematic computations.

4.2 Remote Roomba

During an exchange program at Iowa State University the embedded programming course was heavily reliant on remote sensing and navigation. The challenge was to write drivers in order to integrate an array of sensors and interpret the data delivered from those in near real-time. From this data the development team had to navigate a unknown course, avoiding obstacles and locate the target location.

Even though this challenge is not in traditional sense, computer vision, it does showcase an understanding of data and data analysis, along side the ability to represent data from sensors to humans.

This task was completed at highest grade.

5 Closing remarks

From the time spend within the Phase One MSI program, the theory of reflectance, luminescence and the significance of both filters and the wavelength of light sources, this knowledge can be transferred from heritage applications to more modern application. This could be used to detect otherwise difficult, transparent object.

Furthermore, the extensive knowledge acquired from working with 150 mpix. image pipelines gives perspective into efficiently solving multi-camera system pipelines. It is both a matter of algorithmic efficiency and how to solve a given challenge in practice.

References

[1] Sung Joon Ahn, Wolfgang Rauh, and Sung Kim. "Circular Coded Target for Automation of Optical 3D-Measurement and Camera Calibration." In: *International Journal of Pattern Recognition and Artificial Intelligence* 15 (Sept. 2001), pp. 905–919.

[2] G. Bradski. "The OpenCV Library." In: *Dr. Dobb's Journal of Software Tools* (2000).

[3] OpenCV Contributors. *OpenCV: Calibration and 3D Reconstruction*. 2025. URL: `https://docs.opencv.org/4.x/d9/d0c/group__calib3d.html` (visited on 09/12/2025).

[4] Phase One. *Phase One Calibration Lab*. LinkedIn video post. 2024. URL: `https://www.linkedin.com/posts/phase-one_phaseone-calibrationlab-activity-7172886524177379328-ohD5/` (visited on 09/12/2025).

[5] Phase One. *Heritage Solutions: Multispectral Imaging*. 2024. URL: `https://www.phaseone.com/heritage-solutions/multispectral-imaging/` (visited on 09/12/2025).