Introduction

The report's goal is to document the implementation and the analysis of the two-phase computer vision system for camera calibration and pose estimation. The project goal is to bridge the gap between raw digital imagery and 3D spatial understanding through mathematical modeling and effective implementation. Phase 1's purpose is to establish the foundational camera geometry through calibration, while Phase 2 demonstrates the application by solving the relative pose problem using feature matching. The system achieves this by achieving calibration accuracy with a 0.3652-pixel RMS error and provides pose estimation capabilities. The challenge of computer vision is transforming 2D image data into a good 3D representation of the real world. The project's goal is to address these challenges through a two-phase approach that considers both theoretical and practical implementation. The core objective is to develop a system that not only characterizes the camera's intrinsic properties but also leverages this understanding to solve computer vision problems. The project's importance can extend into robotics, reality systems, and autonomous vehicle perception. Establishing camera geometry precisely and demonstrating its practical application will provide a foundation for computer vision systems that require an accurate understanding of visual data. I selected relative pose estimation as the focus. The selection is driven by compelling factors that are useful in practical computer vision subjects. Relative pose estimation is a building block in computer vision systems, serving as the cornerstone for 3d reconstruction pipelines. This provides richer learning opportunities while demonstrating deeper computer vision concepts.

Phase 1

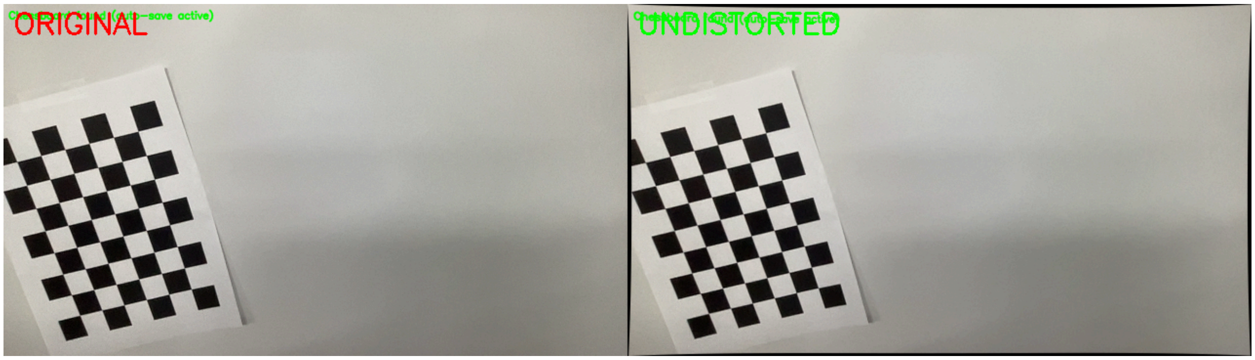
The calibration process utilizes an Apple MacBook Air M1 chip, which incorporates modern computational photography hardware and advanced image processing capabilities. It features a FaceTime HD camera that is 720P in quality with a 720p resolution (1280 x 720 pixels), providing good detail and accurate corner detection. For the calibration pattern, we used a 7 x 6 internal corner chessboard, and provided 54 detectable corners per image. This pattern size is ideal because it strikes an optimal balance between detection reliability and practical usability. The chessboard pattern was printed with great precision to maintain planarity during the image. We successfully captured 34 images using an automated capture script, which exceeds the minimum of 15-20 images specified in the project guidelines. Each image is captured under systematic conditions to ensure robust estimation. The hardware configuration represented an accessible and practical setup, and high-quality camera calibration can be achieved using good consumer equipment with a proper understanding of the methodology. The corner detection phase employed OpenCV's "cv2.findChessboardCorners()" function with careful parameter tuning. Following initial corner detection, we applied pixel refinement using "cv2.cornerSubPix() with a termination criteria of 30 iterations or 0.001 pixel accuracy. This refinement process proved crucial for achieving the exceptional accuracy in the results. Cv2.calibrationCamera() was used with the standard implementation appropriation. The function helped collect 3D-2D points to compute the intrinsic camera matrix. The algorithm solves the following projection. $s * [u\ v\ 1]^T = K\ [R|t]\ [X\ Y\ Z\ 1]^T$. K is the intrinsic camera matrix, Rt denotes the extrinsic parameter for the images, and x,y,z, and u,v represent the world and image coordinates. The CV.2undustort() helps calibration quality too. The calibration achieved an amazing quality with a 0.3652 pixel RMS error rate, placing it in the excellent category by computer vision standards. The intrinsic parameters revealed: Focal lengths $f_x = 985.09$, $f_y = 987.60$ pixels, principal point $c_x = 639.25$, $c_y = 351.52$ pixels, and distortion coefficients $k_1 = 0.068$, $k_2 = -0.185$, $k_3 = 0.052$, $p_1 = 0.003$, $p_2 = 0.001$. The images chessboard_736_comparison.png and 333_comparison.png. Demonstrate the successful distortion correction. Looking at each side-by-side comparison shows how the original images with visible barrel distortion edges are transformed into a geometrically correct representation with perfect straight lines. Examining the undistorted images (chessboard 073 undistorted.png) reveals consistently found auto save messages, which confirm that the corner detection accuracy post-correction is confirmed.

Phase 2

The relative pose estimation problem determines the precise rotation and translation between two camera positions using feature matching and geometric constraints. The approached goal is to leverage epipilar geometry, where F, the fundamental matrix, satisfies the corresponding points. Using the calibrated intrinsic parameters, essential matrix gets computed to enable the recovery of rotation R and translation t though matrix decomposition. The implementation employs a feature detection pipeline using ORB as the main detector with 3000 features per image, palling back to sift when insufficient matches were found. Feature matching uses brute force for ORB and FLANN for sift. The robust estimation utilized the 8 point algorithm with ransac to compute the fundamental matrix, followed by essential matrix computation and pose revecry through svd decomposition. The computer processed image pairs with chessboard_073.png and Chessboard_260.png with using the relative pose with rotation matrix and translation. The visual proof that validated this is good_matches.png and inlier_match.png shows the feature correspondence establishment and outlier rejection. The epipolar_lines.png visualization demonstrate precise alognmwent of the points with computed epiplars lines, which confirms the geometric correctness.

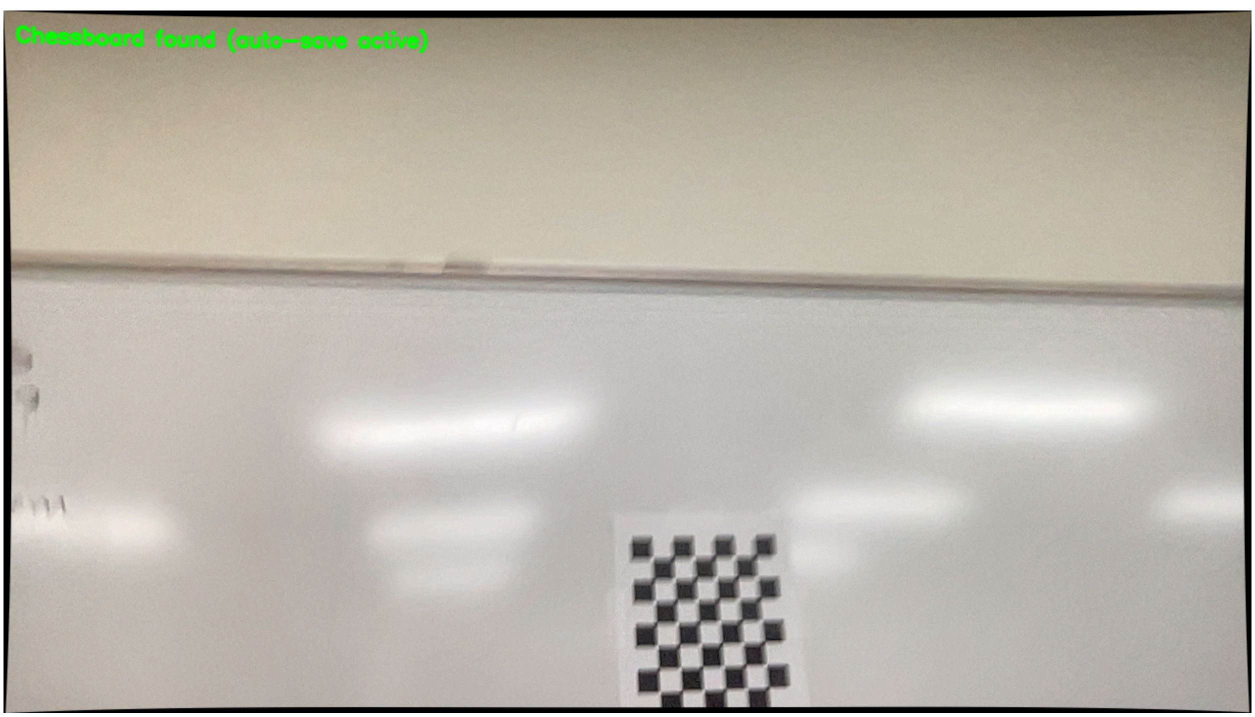
Conclusion

The project demonstrate a complete computer vision understanding from camera characterization to practical application. Phase 1 achieved exceptional calibration accuracy using Macbook air M! Hardware, while phase 2 provide solutions to challenging the pose estimation problems. The integration between phases show how mathematical rigor can be combined with implementation which enables accurate spatial understanding from visual data. The limitation is the scale ambiguity intranslation estimation and the dependence of feature rich environments. The computation requirements are a limitation because it can be reasonable ofr offline processing, but real time application can be limiting. Future improves ca be having a mulit view adjustment for improving the robustness, real time performance optimization, fusion with IMU for metric slace recovery, and deep learning integration for better feature detection. The techniques demonstrated foundation for visual odometry and 3d reconstruction systems, which provided valuable preparation for computer vision challneges.

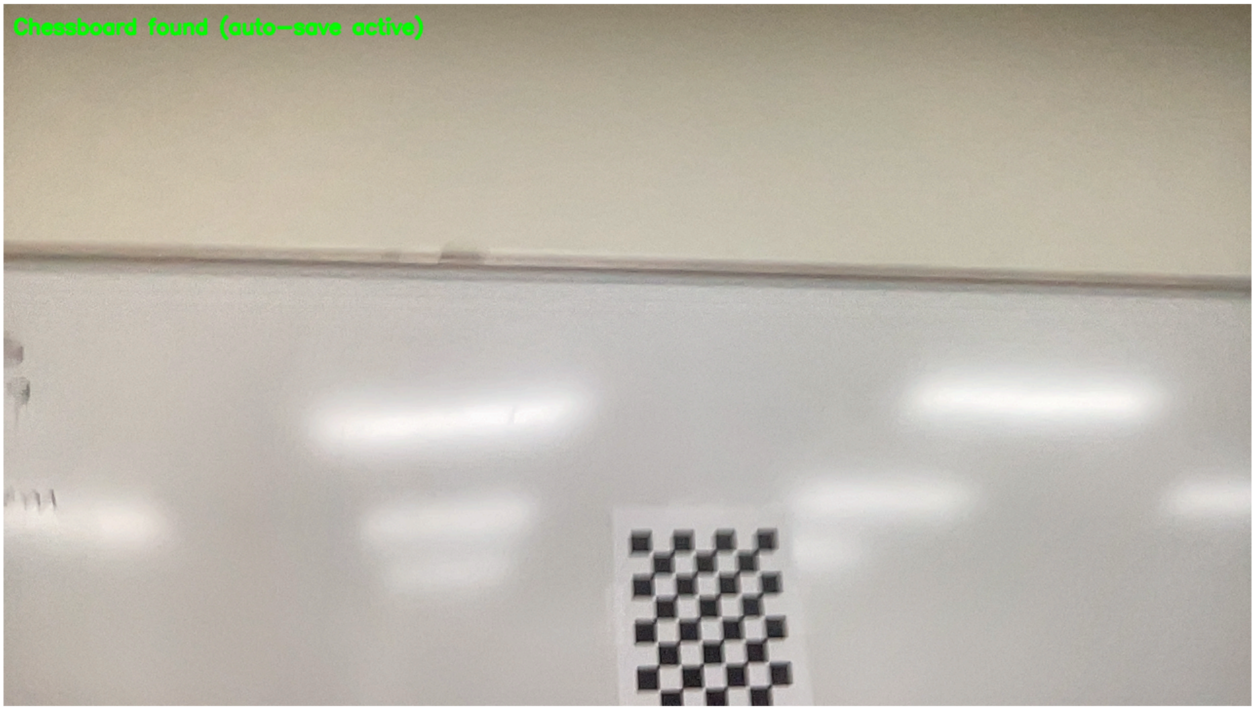


Chessboard_20251106_150459_736_comparison.png

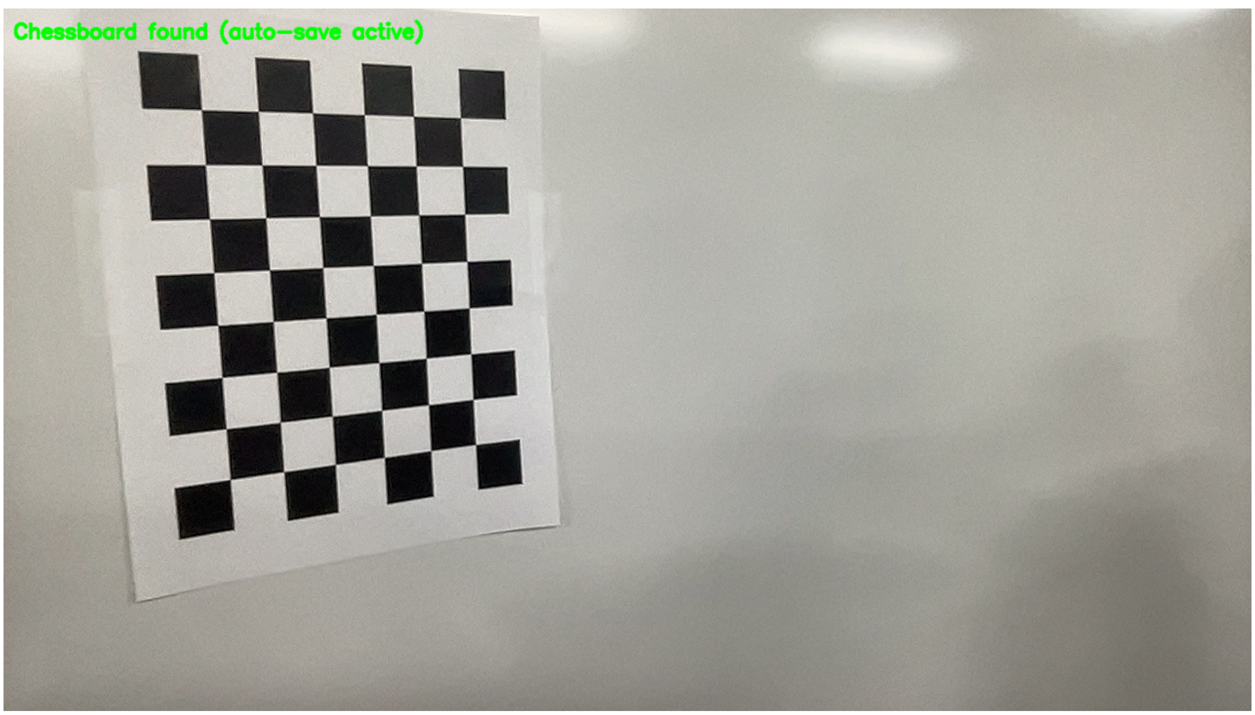
Chessboard_20251106_150545_333_comparison.png.png



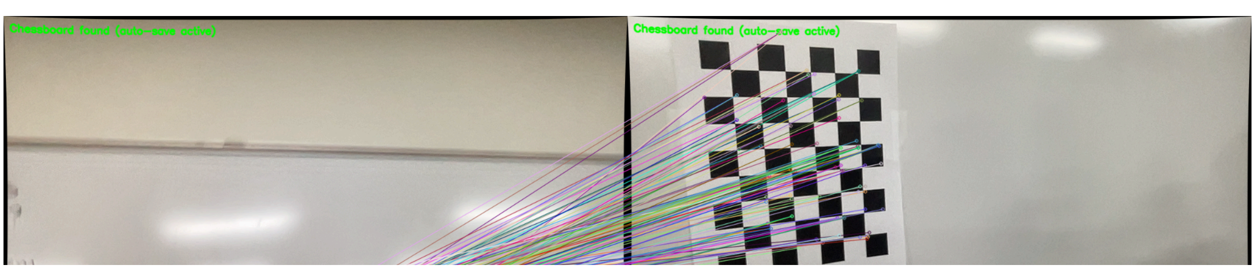
Chessboard_20251106_150440_073_undistorted.png



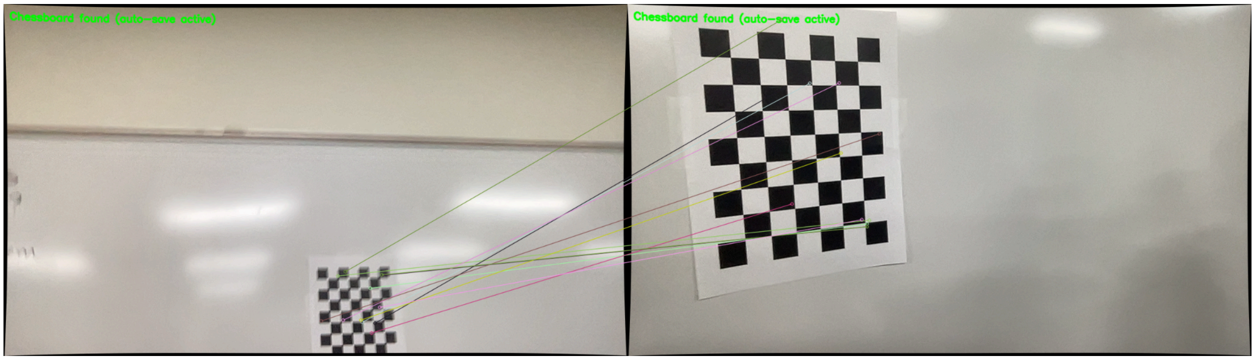
Chessboard_20251106_150440_073.png



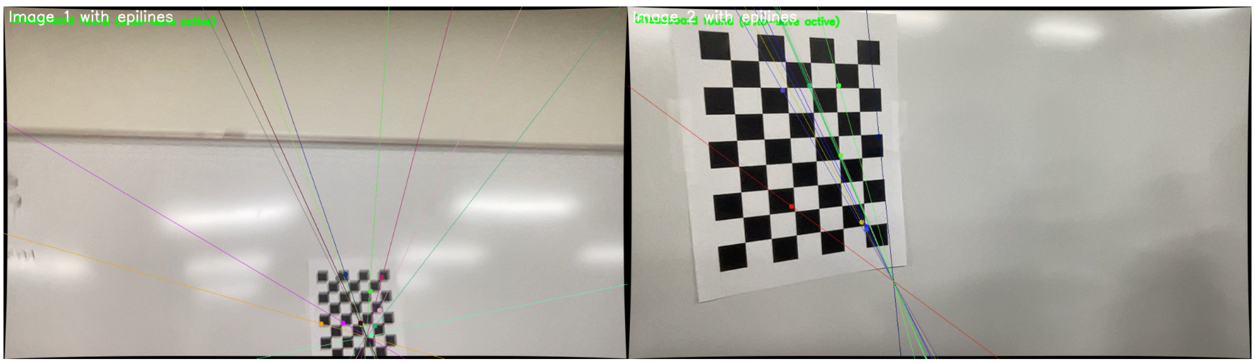
Chessboard_20251106_150445_260.png



Good_matches.png



Inlier_matches,png



Epipolar_lines.png

Work Cited

OpenCV Documentation. (2023). Camera Calibration and 3D Reconstruction