

Beyond Checkerboards: Advantages of Photogrammetric Camera Calibration for Robust 3D Vision and Novel View Generation

Akshay Bharadwaj^{1*} and Derek D Lichti¹

^{1*}Department of Geomatics Engineering, University of Calgary, Calgary, AB, CA.

*Corresponding author(s). E-mail(s): akshay.bharadwaj@ucalgary.ca;
Contributing authors: ddlichti@ucalgary.ca;

Abstract

Imagery is fundamental to modern scientific research, making robust intrinsic camera calibration indispensable for accurate visual inference. The checkerboard-based calibration method has long been favored for its simplicity and ease of deployment, and is widely used even in mission-critical computer vision pipelines. However, its limitations in modeling high-precision camera geometry can compromise downstream performance in tasks requiring geometric accuracy. In this work, we revisit intrinsic calibration through the lens of photogrammetric self-calibration. We show that these alternatives consistently outperform the checkerboard method in both accuracy and precision across a range of vision tasks, including 3D reconstruction, SfM, visual SLAM, and novel view synthesis and reconstruction. Our findings advocate for a paradigm shift toward calibration methods that better reflect the physical and projective properties of camera systems in real-world deployments for critical computer vision applications.

Keywords: Camera Self-Calibration, Bundle Adjustment, Visual Odometry, Structure-from-Motion, NERF

1 Introduction

Cameras have become indispensable sensing modalities in robotics, automation, and computer vision. A vast majority of modern perception pipelines—including reconstruction, localization (visual odometry and simultaneous localization and mapping (SLAM)), object detection, and scene understanding—are fundamentally built upon image data. In safety-critical domains such as autonomous driving and unmanned aerial vehicle (UAV) navigation, cameras serve as primary exteroceptive sensors for collision avoidance and situational awareness. However, raw imagery is not a faithful representation of the 3D world.

Camera systems, especially when paired with complex optics, can introduce radial and decentering (tangential) distortions and chromatic aberration. To enable reliable downstream processing, these distortions must be explicitly modeled and corrected, necessitating accurate camera calibration.

Early calibration strategies were rooted in photogrammetry, where cameras were parameterized using observations of precisely surveyed 3D control targets (Brown 1971), technique still used in close-range photogrammetry, where the high accuracy of reconstruction is needed in object space which translated to a precision of 0.1 px or better in pixel space (Luhmann et al. 2016). Subsequently, self-calibration methods emerged, which

recover both intrinsic and extrinsic parameters by exploiting target correspondences across multiple viewpoints (Maybank and Faugeras 1992; Triggs 1998). Following the maturity of self-calibration methods, closed-form techniques started gaining popularity, most notably Zhang’s method (Zhang 2000), which estimates intrinsics from as few as two images of a planar calibration pattern. While robust and computationally efficient, planar calibration was never designed for high-precision computer vision tasks. Rather, it provides coarse approximations suitable for casual use, compared to metrically accurate calibration required in rigorous 3D measurement and robotics applications.

Despite these limitations, planar calibration has become the de facto standard, aided by its ease of use and implementation in popular toolkits. Variants employing checkerboards, symmetric/asymmetric circle grids, AprilTags (Olson 2011; Richardson et al. 2013), and ArUco/ChArUco (Kedilioglu et al. 2021) boards are now widely adopted. These methods rely on extracting target features (e.g., corners or centers) across multiple images to solve for intrinsics. Alongside this shift, reprojection error has become the dominant criterion for assessing calibration quality. However, reprojection error merely reflects how well the model fits the observed image data; it does not measure the true objective of calibration, to accurately recover 3D structure from 2D imagery. Consequently, calibrations that minimize reprojection error may still produce substantial geometric biases in reconstruction or localization tasks.

Another pervasive issue lies in the parameterization of camera intrinsics. Most widely used computer vision libraries adopt a four-parameter model consisting of focal lengths (f_x, f_y) and principal point (c_x, c_y). This formulation implicitly assumes non-square pixels and enforces independent focal length estimation in x and y , which usually is an over-parameterization and may introduce systematic errors during reconstruction (Fraser 2013). Some studies advocate simplifying the model by enforcing $f_x = f_y$ in downstream tasks, effectively assuming square pixels (Tarrío and Pedre 2015). However, this practice is not broadly recognized in the community, leading to inconsistencies across pipelines (Wang et al. 2020).

In this work, we revisit photogrammetric self-calibration (PSC) and systematically compare it against the widely adopted closed form (CF) planar calibration. Specifically, we calibrate a multi-camera mobile mapping platform using both methods and quantify their impact on downstream tasks requiring geometric accuracy. We further reprocess checkerboard calibration datasets with PSC to evaluate interior orientation parameters (IOPs) within a controlled sparse 3D reconstruction against surveyed ground truth. Finally, we extend our analysis to benchmark datasets, including EuRoC-MAV (Burri et al. 2016) and Oxford-spires (Tao et al. 2025), to evaluate improvements in visual odometry accuracy.

Our experiments demonstrate that photogrammetric self-calibration consistently yields superior performance in tasks demanding precise geometric fidelity. While CF planar calibration remains useful for quick approximations and non-critical computer vision tasks, it is fundamentally inadequate for high-accuracy applications such as dense reconstruction, SLAM, and benchmarking visual systems. We argue for the adoption of PSC in widely used benchmarks (e.g., KITTI (Geiger et al. 2013), TUM-VI (Schubert et al. 2018), Waymo (Sun et al. 2020)), where precise sensor modeling is critical for fair evaluation of visual algorithms.

Our contributions in this paper are,

- We calibrate a multi-camera system using both photogrammetric self-calibration and checkerboard-based closed-form methods, and analyze their impact on downstream tasks.
- We systematically compare PSC and CF calibration methods
- We apply photogrammetric self-calibration to checkerboard calibration datasets and compare intrinsic orientation parameters via sparse 3D reconstruction against surveyed control points.
- We extend photogrammetric self-calibration to benchmark datasets (EuRoC-MAV, Oxford-spires) and demonstrate improved accuracy in visual odometry.

2 Revisiting Calibration

2.0.1 Closed Form Calibration

Since the CF calibration method is widely used, let us revisit the CF calibration method proposed by (Zhang 2000) and contrast it with the PSC method presented by (Lichti et al. 2020). In the CF calibration, the calibration relationship is given by (1), where $\mathbf{K}$ is the camera matrix (referred to as IOPs in photogrammetry), $[R|t]$ represents the extrinsic parameters from the world frame to the camera frame, $\tilde{P} = [X, Y, Z, 1]^T$ is a 3D point in homogeneous coordinates, $\tilde{p} = [x_0, y_0, 1]^T$ is the 2D point on the image plane, and s is an arbitrary scale factor. The camera matrix $\mathbf{K}$ defines how the 3D point in the camera frame projects onto the image plane, where c_x, c_y are the principal distances (focal lengths) in x and y , respectively, x_p, y_p are the principal point coordinates, and γ is the skew factor between the two image axes.

$$s\tilde{p} = \mathbf{K}[R|t]\tilde{P} \quad \text{where} \quad \mathbf{K} = \begin{bmatrix} c_x & \gamma & x_p \\ 0 & c_y & y_p \\ 0 & 0 & 1 \end{bmatrix} \quad (1)$$

The intuition behind the (Zhang 2000) method is to observe a minimum of four points on a plane from at least two poses, achieved either by moving the camera or the model plane is to estimate the exterior orientation parameters (EOPs) in an initial step, and subsequently the IOPs and lens distortion coefficients in distinct steps. The method assumes the model plane lies at $Z = 0$, which simplifies (1) to $s\tilde{p} = \mathbf{H}\tilde{P}$ where $\mathbf{H} = \mathbf{K}[r_1, r_2, t]$ and $\mathbf{R} = [r_1, r_2, r_3]$. (Zhang 2000) utilizes the image of the absolute conic, defined by $\mathbf{K}^{-T}\mathbf{K}^{-1}$, to infer constraints (2) and (3) where $\mathbf{H} = [h_1, h_2, h_3]$, to arrive at a CF solution.

$$h_1^T \mathbf{K}^{-T} \mathbf{K}^{-1} h_2 = 0 \quad (2)$$

$$h_1^T \mathbf{K}^{-T} \mathbf{K}^{-1} h_1 = h_2^T \mathbf{K}^{-T} \mathbf{K}^{-1} h_2 \quad (3)$$

For the CF solution, the image of the absolute conic is described by $\mathbf{M} = \mathbf{K}^{-T} \mathbf{K}^{-1}$, a symmetric matrix that can be vectorized into a 6D vector $m = [M_{11}, M_{12}, M_{22}, M_{13}, M_{23}, M_{33}]$. Using constraints from (2) and (3), a solvable linear homogeneous system is derived based on observations from image data, as shown in (4), where v_{ij}

encapsulates the products and sums of h_i and h_j from (2). For a calibration procedure, if n such images of model planes are observed, stacking n instances of (4) yields (5). Once solved for m , estimating the IOPs becomes straightforward. For more details on estimating the IOPs and homography, refer to (Zhang 2000). At least two images of the model plane are necessary to estimate principal distance and principal point (assuming zero skew), whereas more than three images are needed to estimate the complete set of IOPs.

$$[v_{12}^T (v_{11} - v_{22})^T] m = 0 \quad (4)$$

$$\mathbf{V}m = 0 \quad (5)$$

It is trivial to determine all the intrinsics once $\mathbf{m}$ is determined. This initial estimate is further refined using maximum likelihood estimation by minimizing (6),

$$\sum_{i=1}^n \sum_{j=1}^m |\mathbf{p}_{ij} - \hat{\mathbf{p}}(\mathbf{K}, R_i, t_i, \mathbf{P}_j)|^2 \quad (6)$$

2.0.2 Photogrammetric Self-Calibration

The PSC approach is modeled by the collinearity equations given by (7) and (8), which show the theoretical straight-line relationship between the point $\mathbf{P}_i$, its projection on the image $\mathbf{p}_i$, and the perspective center of the image $\mathbf{P}^c(X^c, Y^c, Z^c)$. The principal point is denoted by $[x_p, y_p]^T$, and the elements of the rotation matrix $\mathbf{R}$ are represented by r_{pq} .

$$\begin{aligned} x_{ij} + \epsilon_{x_{ij}} &= x_{p_j} - \\ & c_j \frac{r_{11}(X_i - X_j^c) + r_{12}(Y_i - Y_j^c) + r_{13}(Z_i - Z_j^c)}{r_{31}(X_i - X_j^c) + r_{32}(Y_i - Y_j^c) + r_{33}(Z_i - Z_j^c)} + \Delta x_{ij} \end{aligned} \quad (7)$$

$$\begin{aligned} y_{ij} + \epsilon_{y_{ij}} &= y_{p_j} - \\ & c_j \frac{r_{11}(X_i - X_j^c) + r_{12}(Y_i - Y_j^c) + r_{13}(Z_i - Z_j^c)}{r_{31}(X_i - X_j^c) + r_{32}(Y_i - Y_j^c) + r_{33}(Z_i - Z_j^c)} + \Delta y_{ij} \end{aligned} \quad (8)$$

The Brown–Conrady distortion model (Brown 1971), shown in (9) and (10), is used to model lens distortion. The effects of lens distortion, given by

$[\Delta x, \Delta y]^T$, are added to the collinearity equations (7) and (8). Expression (9) describes radial distortion, with k_i as the coefficients, while (10) models decentering distortion, where p_1 and p_2 are the decentering distortion coefficients. r is the radial distance with respect to the principal point $[x_p, y_p]^T$ and $\bar{x}, \bar{y}$ are image coordinates reduced by the principal point.

$$\begin{aligned}\Delta x &= \bar{x}(k_1 r^2 + k_2 r^4 + \dots + k_n r^{2n}) \\ \Delta y &= \bar{y}(k_1 r^2 + k_2 r^4 + \dots + k_n r^{2n})\end{aligned}\quad (9)$$

$$\begin{aligned}\Delta x &= p_1(r^2 + 2)\bar{x}^2 + 2p_2\bar{x}\bar{y} \\ \Delta y &= p_2(r^2 + 2)\bar{x}^2 + 2p_1\bar{x}\bar{y}\end{aligned}\quad (10)$$

The IOPs, EOPs, and lens distortion terms are estimated simultaneously using self-calibrating bundle adjustment (BA), which is formulated as a constrained Gauss–Markov model. A general linearized form of self-calibrating BA is given in (11), where the design matrix is parametrized into IOPs ($\mathbf{A}_i$), EOPs ($\mathbf{A}_e$) and object point coordinates ($\mathbf{A}_o$), and $\hat{\delta}_e$, $\hat{\delta}_i$, $\hat{\delta}_o$ are the correction vectors to the initial values, respectively; $\mathbf{w}_p$ is the misclosure vector, and $\mathbf{v}_p$ is the residual vector.

$$\mathbf{A}_e \hat{\delta}_e + \mathbf{A}_i \hat{\delta}_i + \mathbf{A}_o \hat{\delta}_o + \mathbf{w}_p = \hat{\mathbf{v}}_p \quad (11)$$

Further inner constraints can be imposed on object points, as defined by the network datum in (12), where $\mathbf{G}_o$ is the design matrix (Luhmann et al. 2014). The inner constraints help remove the rank deficiency arising from inherent undefined datum (gauge) elements.

$$\mathbf{G}_o^T \hat{\delta}_o = 0 \quad (12)$$

2.0.3 Comparison

PSC and CF calibration offer fundamentally different approaches for camera calibration. In PSC, all parameters, IOPs, lens distortion coefficients, EOPs, and target coordinates, are jointly optimized in a single BA, enabling robust modeling but requiring careful network design. Without sufficiently diverse camera poses and robust scene geometry, PSC can fail to converge. CF calibration, in contrast, separates estimation into sequential steps: extrinsic parameters are determined

first using precise calibration targets, followed by computation of camera intrinsics and distortion coefficients, yielding a CF solution whenever the calibration is performed. While CF methods always produce solutions that try to fit the calibration observations, they often absorb systematic errors, such as lens distortion or target misalignment, into the intrinsic parameters (Fraser 2013), making these biases difficult to disentangle or even detect. Since CF method is closed form, very little attention is given towards photogrammetric network design, which can produce a highly biased result. In terms of evaluation, CF calibration typically relies on reprojection error, which measures the consistency of image point projections with model predictions, while PSC additionally considers reconstruction error, quantifying how accurately the scene geometry itself is recovered along with being robust to outlier observations. Overall, CF calibration provides guaranteed results within controlled settings but may mask systematic errors, whereas PSC is more flexible and potentially offers increased accuracy across more challenging scenarios, as long as the calibration network is well designed to avoid degeneracies.

2.1 Camera Calibration

In this subsection, we discuss the calibration of the IOPs and distortion coefficients of a single camera within a multi-sensor mobile mapping platform comprised of four cameras. The cameras were calibrated in the camera calibration laboratory at the University of Calgary, which features hundreds of calibration targets distributed around the walls and ceilings (Figure 1). The 3D coordinates of the target centers were accurately measured using a terrestrial LiDAR scanner.



Fig. 1: Calibration laboratory used for calibration of camera intrinsics

For each camera, the target field was imaged from four different locations in the laboratory and at various azimuth angles ensuring convergent imagery according to good first order network design. At each azimuth angle, two images were captured with a 90° change in the rotation angle to help decouple the x and y principal point parameters (Vo et al. 2021). The target centers were extracted and labeled in the images for use in the PSC BA. The known 3D coordinates of the targets served as control points in the BA, enabling estimation of reconstruction errors for these control points, but never treated as fixed in PSC.

2.1.1 Model identification

A key consideration during camera calibration is the selection of an appropriate model. While the choice of the camera projection model is typically straightforward, selecting a suitable distortion model is less clear-cut. The literature most commonly references the radial-decentering (Brown–Conrady) model (Brown 1971) and the equidistant model (Kannala and Brandt 2006). The former accounts for both radial and decentering distortions, whereas the latter includes only radial terms.

To choose the optimal model, and to determine if all lens distortion has been compensated, we follow an iterative model identification approach as suggested in (Lichti et al. 2020). We begin by calibrating the intrinsics using an initial model with a subset of distortion parameters, then plot the radial distance of the keypoints against the radial component of the residuals. If systematic errors are observed in the residuals, additional distortion terms are incorporated into the model. This process is repeated—adding terms and, if necessary, revising the camera model itself—until all systematic errors are eliminated from the residual plots.

3 Experiments and Discussion

This section is organized into 4 sections where we start by discussing the model selection for camera calibration and then demonstrate the superiority of PSC over CF solution in various tasks of computer vision, namely SfM reconstruction,

visual odometry (VO) and novel view synthesis and reconstruction with neural radiance fields (NERF).

3.1 Camera Calibration

3.1.1 Camera

For these experiments, we used a FLIR area scan camera equipped with a CMOS sensor and a global shutter. The camera has a resolution of 3.2 megapixels (2048×1536 pixels) with square pixels of size $3.4\mu\text{m}$. Images were acquired at 10 Hz, triggered by the azimuth angle of the LiDAR sweeps which is determined by the extrinsic calibration between the camera and the LiDAR. The camera was fitted with a 4 mm F/1.8 fixed-focus lens with a fixed aperture. Exposure and gain settings were kept constant during the capture of each data sequence.

3.1.2 Calibration

As discussed in 2.1, the camera was calibrated using PSC in the calibration lab. Images of the calibration target field were captured from four different positions in the lab with convergent imaging geometry. Approximate EOPs were estimated at the time of each image capture for use in the BA. Care was taken to ensure that the calibration targets filled the entire sensor format, and that the same targets were captured at two different camera rotation angles separated by 90° (analogous to portrait and landscape modes).

For calibration, we employed a pinhole projection model with radial-decentering distortion (Section 2.1), initialized it with three radial and two decentering coefficients. After executing the BA, we plotted the radial component of the residuals against radial distance and inspected the plot for systematic errors (Figure 2a). Due to the presence of significant systematic errors in Figure 2a, we repeated the process by adding five radial distortion terms (Figure 2b), and subsequently six (Figure 2c). Since a small amount of systematic error persisted in Figure 2c, we added a seventh radial distortion term (k_7) until all systematic errors were eliminated.

For comparison, we also calibrated the same camera using a checkerboard target and the CF method presented in (Furgale et al. 2013), employing the pinhole camera model with equidistant

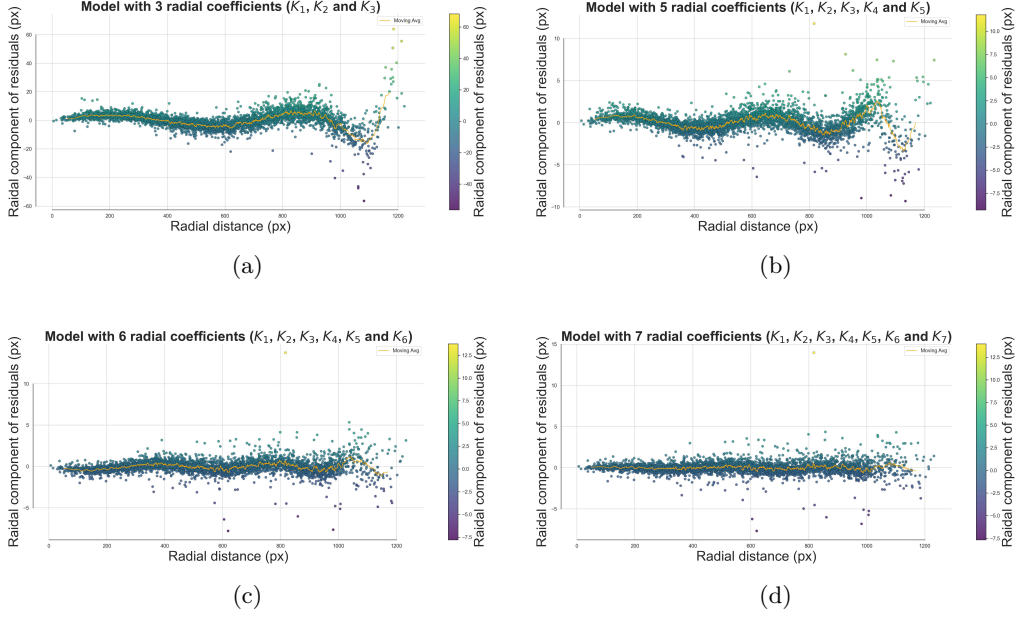


Fig. 2: The figure shows the systematic errors inherent in the image observation points that are due to the radial distortion. Figures (a) thru (d) shows the residual plots when additional radial distortion terms are added to the bundle adjustment

distortion and refined until the reprojection errors were 0.34 px. This approach was chosen as the equidistant model (Kannala and Brandt 2006) is utilized in several benchmark datasets, including (Burri et al. 2016; Schubert et al. 2018; Ramezani et al. 2020; Tao et al. 2025), among others.

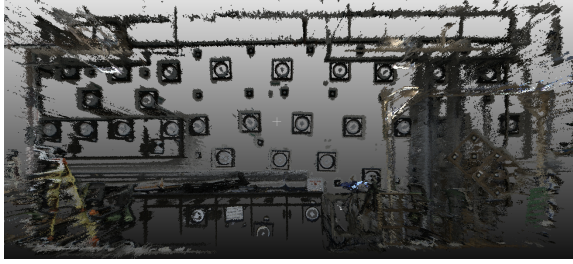
3.2 SfM Reconstruction

SfM reconstruction was performed using data captured in two scenes: one in the calibration laboratory and the other at a larger scale around the Calgary Center of Innovation and Technology (CCIT) building at the University of Calgary. Debayered images were continuously captured at 10 Hz using robotic operating system (ROS) drivers. These images were then imported into COLMAP (Schönberger and Frahm 2016; Schönberger et al. 2016b) for feature extraction (Lowe 2004) and feature matching (Schönberger et al. 2016a). The camera model and intrinsic parameters were provided during the feature extraction and feature matching.

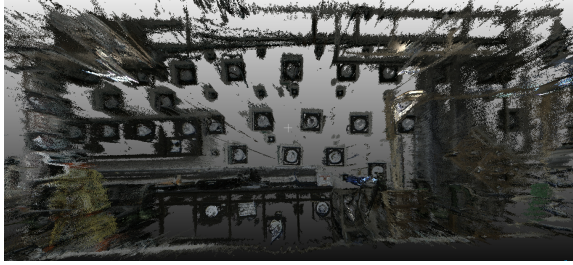
The preprocessed data was subsequently used with GloMap (Pan et al. 2024) to perform global bundle adjustment (BA) and generate a

sparse model, taking care not to refine the camera intrinsics or distortion parameters. COLMAP (Schönberger and Frahm 2016; Schönberger et al. 2016b) was then employed again to densify the sparse pointcloud (PC) into a dense PC. LiDAR PCs, along with inertial information, were processed using FAST-LIVO (Zheng et al. 2025) to generate accurate, scale-sized 3D models of the scenes; these were used to scale the dense reconstructed PCs. The dense reconstructed PC of the calibration laboratory is shown in Figure 3, and that of the CCIT building is shown in Figure 4.

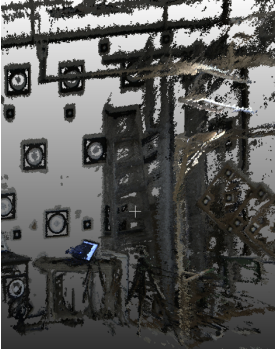
To demonstrate the reconstruction quality, we present both quantitative and qualitative evidence. For the quantitative evaluation, the centers of all calibration targets were measured using a terrestrial LiDAR scanner and the Euclidean distances between all pairs of calibration target centers were computed using these surveyed coordinates to establish ground truth values. Subsequently, the centers of the calibration targets were detected in the reconstructed point clouds, and the Euclidean distances between each pair were similarly calculated. The distance errors were then quantified and reported as the mean absolute



(a)



(b)



(c)



(d)

Fig. 3: Qualitative results from Sfm using COLMAP in "Calibration Lab" dataset.

error (MAE), root mean square error (RMSE), and standard deviation, as summarized in Table 1.

The results in Table 1 clearly demonstrate the superior performance of the PSC method compared to the CF calibration solution. Across all error metrics, PSC yields substantially lower reconstruction errors. Specifically, the MAE is reduced from 0.209 m with CF calibration to 0.038 m with PSC, representing almost an order of magnitude improvement. Similarly, the RMSE decreases from 0.236 m to 0.043 m, indicating that PSC not only reduces the average error but also

Parameter	Closed	Photogrammetric
Method	Form	Self-Calibration
MAE (m)	0.209	0.038
RMSE(m)	0.236	0.043
Std.(m)	0.110	0.020

Table 1: The table show the comparison of distance errors calculated between the reconstructed target centers with that of models estimated with closed form calibration and BA parameters.

mitigates the impact of larger deviations. Furthermore, the standard deviation drops from 0.110 m to 0.020 m, showing that PSC produces 80% more accurate and precise estimates of target positions.

Overall, these results demonstrate that jointly estimating the IOPs, EOPs, lens distortion coefficients, and object point 3D coordinates, combined with a superior geometric network design with PSC, yields a more accurate and reliable 3D reconstruction than the closed-form approach. Such improvements are especially crucial for downstream applications where both high accuracy and consistency of 3D measurements are essential.

This notable improvement indicates that the PSC approach effectively minimizes geometric errors in the reconstruction, yielding both higher accuracy (as reflected in the lower MAE and RMSE) and greater precision (as indicated by the reduced standard deviation). These results empirically validate the use of PSC in the estimation of camera IOPs, lens distortion coefficients and target positions, particularly in Sfm pipelines where accurate spatial relationships are critical. A qualitative comparison of Figures 3a, 3b, 3c, and 3d further illustrates this behavior, especially when inspecting points around the calibration targets. From these figures, it is evident that the PCs obtained via CF calibration method are noticeably noisier than those generated using PSC. This increased noise arises from residual systematic errors in the estimated IOPs and distortion coefficients, which fail to fully correct for lens distortion.

These errors become even more pronounced when reconstruction is performed on larger-scale

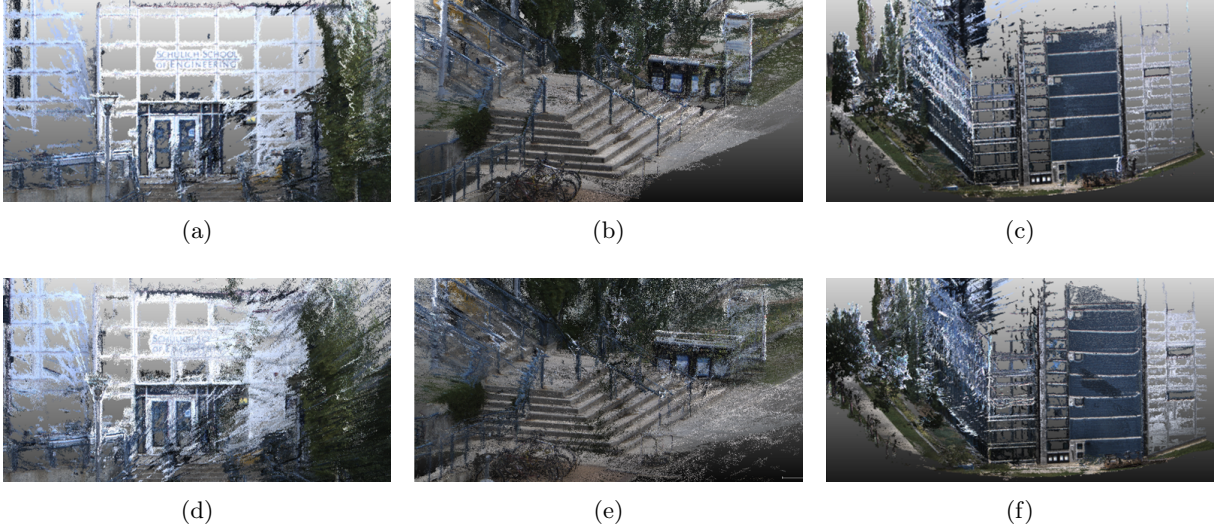


Fig. 4: The figure shows the qualitative comparison of photogrammetric reconstruction of an engineering building in University of Calgary. Figures a-c shows various sections of the reconstruction performed using the BA parameters and figures d-e shows the same section when reconstructed using closed form parameters

scenes, such as that shown in Figure 4. For example, when comparing Figures 4a and 4d, the building name is clearly legible in 4a but blurred in 4d. Similarly, the geometry of bicycle tires in Figure 4b is more accurately preserved compared to the distorted shapes observed in Figure 4e. Another striking example is observed when comparing Figures 4c and 4f, where systematic errors manifest as curved building façade lines in the CF reconstruction. The effect intensifies with distance from the camera: near the base, the lines appear straight, but they increasingly curve with height—an artifact that is eliminated with the PSC calibration.

3.3 Visual Odometry

To evaluate the effects of calibration on VO, we collected nine datasets across the University of Calgary campus under diverse lighting conditions, both indoors and outdoors. The spatial extent of the sequences varied from less than 10 m for few indoor captures to approximately 120 m in large-scale outdoor environments. Each sequence was recorded with synchronized LiDAR and inertial data, which were processed using FAST-LIO2 (Xu et al. 2022) to generate maps and trajectories.

These trajectories served as ground truth (GT) for evaluating the VO results.

The captured sequences represent a variety of test scenarios, summarized in Table 2. The sequences are captured indoors and outdoors ranging from small area around a mural to a very large area around the campus library. The diverse range of sequences were captured to test VO performance in diverse conditions. Several of these datasets are also used in Section 3.4 to evaluate the performance of novel view synthesis.

Table 3 presents the RMSE and standard deviation of the ATE obtained when running monocular ORB-SLAM3 (Campos et al. 2021) using camera parameters derived from CF calibration and BA-based PSC which are evaluated using EVO python library (Grupp 2017). The results presented is a summary of running ORB-SLAM3 10 times for each data as presented in (Campos et al. 2021). The results confirm that PSC-based parameters consistently provide superior VO/SLAM performance across nearly all datasets, particularly in large-scale or geometrically challenging environments.

For the long outdoor trajectories, such as "Library Outdoor" and "Kinesiology Loop" (loop closure attempt, handheld), the CF calibration yields very high RMSE values (16.72 m and 6.91

Seq. Name	Mount type	Location	Details	Area (sq m)
Chem-math Hallway	Cart	Indoor	Captured in the chemistry and mathematics building hallway, looping around the stairs for loop closure	≈ 2700
Dino	Handheld	outdoor	Captured around the dino in close range at various altitudes in convergent geometry	≈ 23
Eleutheria	Handheld	indoor	Similar to dino, captured in close range around Eleutheria statue	≈ 24
Kinesiology loop	Handheld	outdoor	Data was captured outdoor while walking around open space with varying height and looped around	≈ 2400
Library outdoor	Cart	outdoor	Captured outdoor while pushing cart.	≈ 5500
Mural	Handheld	indoor	Very close range capture with rapid movements with convergent imagery and varying height	≈ 7.5
Olympic arch	Cart	outdoor	Captured outdoor while looping olympic arch on campus	≈ 400
Science-B hallway	Cart	Indoor	Captured indoor similar to Chem-Math hallway	≈ 1600
Ucalgary Motto	Cart	Outdoor	Captured outdoor on cart around the motto structure in an convergent geometry	≈ 168

Table 2: Summarized details of the data collected by the multi-camera system for evaluating visual odometry

m, respectively), indicating severe scale drift and poor trajectory consistency. In contrast, the PSC calibration dramatically reduces the RMSE to 1.21 m and 0.28 m, respectively, representing improvements of more than an order of magnitude. These results suggest that PSC calibration greatly enhances metric consistency in large outdoor environments and improves robustness in trajectories requiring loop closure.

Similarly, for the medium-scale indoor cart-mounted captures, such as "*Chem-Math Hallway*" and "*Science-B Hallway*", PSC calibration again yields significantly lower RMSE values (0.43 m and 0.52 m) compared to CF calibration (3.73 m and 4.39 m). This demonstrates that correcting residual distortion effects through PSC calibration reduces cumulative drift and produces trajectories

that align more consistently with ground truth over longer indoor corridors.

For the shorter or smaller-scale sequences, such as "*Mural*", "*Olympic Arch*", and "*Ucalgary Motto*", both methods achieve lower absolute errors. Nevertheless, PSC calibration further refines accuracy; for example, in "*Olympic Arch*" the RMSE improves from 0.39 m to 0.19 m, and in "*Ucalgary Motto*" from 0.25 m to 0.09 m. While the relative benefit is smaller in constrained scenes, these improvements highlight that BA calibration consistently minimizes subtle biases that accumulate even over short distances.

One exception is the "*Dino*" dataset, where both methods achieve comparable performance (≈ 0.15 m RMSE). This can likely be attributed

		CF	PSC
Chem-math hallway	RMSE(m)	3.736	0.430
	Std.(m)	0.108	0.014
Dino	RMSE(m)	0.154	0.155
	Std.(m)	0.006	0.001
Eleutheria	RMSE(m)	0.103	0.063
	Std.(m)	0.009	0.013
Kinesiology loop	RMSE(m)	6.916	0.284
	Std.(m)	6.604	0.151
Library outdoor	RMSE(m)	16.724	1.217
	Std.(m)	2.446	0.205
Mural	RMSE(m)	0.084	0.064
	Std.(m)	0.005	0.003
Olympic arch	RMSE(m)	0.395	0.188
	Std.(m)	0.217	0.009
Science-B hallway	RMSE(m)	4.399	0.522
	Std.(m)	0.094	0.074
Ucalgary motto	RMSE(m)	0.246	0.097
	Std.(m)	0.008	0.001

Table 3: Transposed comparison of RMSE absolute trajectory error (ATE) running Monocular Visual SLAM using ORB-SLAM3 (Campos et al. 2021) with parameters estimated using CF based calibration method and PSC parameters determined by running PSC on the same data.

to its small capture scale and limited scene variation, where the intrinsic distortions have relatively minor impact on VO performance.

In terms of precision, as indicated by the standard deviations, PSC calibration not only reduces mean error but also results in far more stable trajectory estimates across all sequences. For instance, in Kinesiology Loop, the trajectory variance reduces from 6.60 m with CF calibration to just 0.15 m with PSC calibration, reinforcing that PSC calibration substantially stabilizes VO across challenging motion paths.

Overall, these results confirm that BA-based PSC calibration provides trajectories that are both more accurate and more consistent, particularly in long-range outdoor captures and extended indoor corridor sequences where systematic distortion errors in CF calibration would otherwise accumulate. This demonstrates the practical

importance of PSC in improving VO/SLAM reliability, particularly when applied to large-scale mapping and navigation tasks.

3.4 Novel View Synthesis and Reconstruction

Novel view synthesis and reconstruction has emerged as a state-of-the-art approach in computer vision, enabling the generation of new viewpoints of a scene from a limited set of input images using methods such as NERF and 3D Gaussian splatting. NERF, in particular, has gained tremendous popularity for its ability to produce photorealistic renderings and achieve remarkable 3D scene modeling. The rapid growth of research in this area motivates an investigation into the effects of IOPs on the training of NERF models.

To this end, we use three sequences from the previously discussed odometry dataset, of visually distinct landmarks around University of Calgary campus, in addition to the "Calibration Lab" sequence previously used in SfM experiments. Consistent with the SfM pipeline, the captured data was pre-processed with COLMAP to estimate the approximate camera poses and construct sparse models, which were subsequently imported into NERF training. For training, we employed Nerfstudio's Nerfacto network model (Tancik et al. 2023), with each sequence trained for 200K steps using the default settings. Care was taken to avoid optimizing the IOPs during both preprocessing and training. Since Nerfstudio does not currently support the Full-OpenCV camera model, all images were rectified prior to COLMAP preprocessing to ensure compatibility with the training pipeline.

For evaluation, 10% of the input images were randomly selected from the sequence. The results of the quantitative evaluation on these images are presented in Table 4, while Figure 5 provides a qualitative and visual comparison of the reconstructed renderings. The Figure 5 shows four scenes, presented in a tabulated layout, where the first two rows present results from "calibration lab", the next two for "Dino" then "Eleutheria" and "Olympic arch". The 1st two columns show different rendering, where as the third column shows the generated PC.

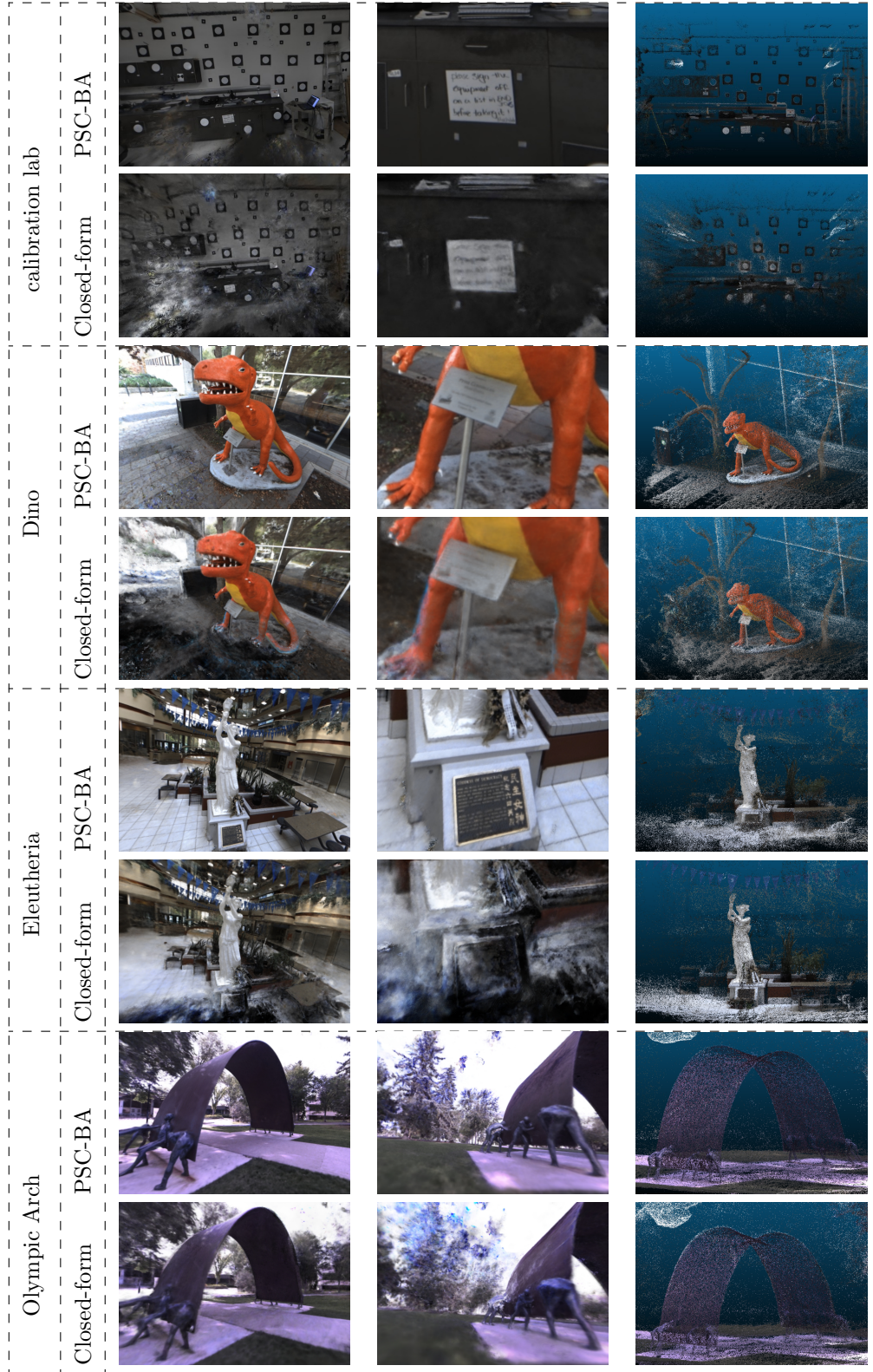


Fig. 5: Qualitative results of novel-view reconstruction using NERF (Tancik et al. 2023) for four different scenes around the University of Calgary campus. The first two columns show renderings from two different scenes, where as the third column shows the generated PC

	Calibration Lab		Dino		Olympic Arch		Eleutheria	
	PSC	CF	PSC	CF	PSC	CF	PSC	CF
PSNR $\uparrow$ (mean)	25.460	20.023	23.197	16.407	24.106	17.674	24.012	17.50
(std)	1.329	1.819	2.770	1.931	1.930	1.466	2.388	2.482
SSIM $\uparrow$ (mean)	0.914	0.705	0.861	0.568	0.807	0.634	0.913	0.693
(std)	0.032	0.064	0.031	0.060	0.042	0.045	0.013	0.052
LPIPS $\downarrow$ (mean)	0.127	0.411	0.129	0.396	0.275	0.466	0.089	0.341
(std)	0.023	0.074	0.027	0.038	0.046	0.044	0.015	0.044

Table 4: Quantitative comparison of NeRF reconstruction using imagery with camera intrinsics determined using photogrammetric PSC and CF parameters across different datasets.

The quantitative evaluation presented in Table 4 compares the performance of NeRF reconstructions using camera intrinsics estimated via PSC and a CF solution.

Across all datasets and metrics, the BA-derived IOPs and distortion coefficients consistently outperform the CF parameters. This trend is most evident in peak signal to noise ratio (PSNR) and structural similarity index measure (SSIM) scores, both of which are higher for the PSC parameters, indicating better image fidelity and structural similarity in the synthesized views. In contrast, learned perceptual image patch similarity (LPIPS), which reflects perceptual similarity (lower is better), is consistently lower for the PSC parameters, further corroborating the improved visual quality of reconstructions.

"*Calibration Lab*" which is a controlled environment with calibration targets shows the highest PSNR (25.46) and SSIM (0.91) for PSC parameters with the gap between PSC and CF results being substantial, indicating that precise IOPs are critical. Similar performance can be observed across the sequences for "*Dino*" to have the least PSNR among the group for CF parameters. This suggests that inaccurate intrinsics can severely impair reconstruction in scenes with fine detail and complex contours—especially when those details contribute significantly to photometric optimization.

The results in the Table 4 for PSNR shows a consistent advantage (5–7 dB) for PSC parameters across all scenes, indicating that reconstruction accuracy is highly sensitive to the quality of

camera IOPs. SSIM follows a similar trend, though with slightly smaller absolute margins, suggesting that the structural fidelity is better preserved with PSC IOPs. LPIPS tell that reconstructions with CF parameters produce notably more perceptual artifacts, with glslpips values 2–4x worse in some cases.

Across all metrics, the standard deviations are lower for the PSC parameters, implying greater stability and predictability in reconstruction quality. This is particularly important for applications where consistent view synthesis quality is required across frames (e.g., video, simulation, or AR/VR).

The qualitative results in Figure 5 supports the results observed in Table 4. Across the board, reconstruction with PSC parameters show significantly less artifacts after training for 200K iteration as compared to the CF parameters. For example, in the "*Calibration Lab*" rendering with CF parameters a lot more clouding is seen that with the PSC parameters. Comparing the rendering with text in "*Calibration Lab*", "*Dino*" and "*Eleutheria*", we can observe that the rendering with PSC parameters are lot sharper than the rendering with CF parameters. Even the reconstructed PC shows less noise with the reconstruction that with CF IOPs. Observing the rendering from "*Calibration Lab*", we can observe the pipes are curved in the rendering with CF stating that intrinsics are not completely removed by the CF calibration.

3.5 Visual Odometry - Benchmark datasets

In this section we go a step further and apply the PSC to benchmark datasets like EuRoC-MAV (Burri et al. 2016) and Oxford-spires (Tao et al. 2025). The goal is to show that the PSC can be used to improve the VO performance of state-of-the-art visual SLAM systems like ORB-SLAM3 (Campos et al. 2021).

In order to achieve that, we first choose benchmark datasets that provide raw calibration data and distinctive features to add variety to the analysis. As mentioned before, here we choose EuroC-MAV (Burri et al. 2016) for its popularity and Oxford-spires (Tao et al. 2025) for its diversity in data.

Before running PSC on the calibration data of the benchmark dataset, we preprocess the calibration pattern data with CF calibration toolbox like (The MathWorks 2025; Rehder et al. 2016) to get an estimation of the IOPs, EOPs, image points and control points. With preprocessing, we can now apply the same calibration procedure explained in 2.1 to estimate the IOPs, EOPs and distortion coefficients. The obtained IOPs and the distortion coefficients are then used with ORBSLAM3 (Campos et al. 2021), which is compared to the output with the default parameters provided by the benchmark datasets.

3.5.1 EuroC-MAV Dataset

In this subsection we discuss and compare the RMSE ATE of running monocular visual SLAM and monocular visual-inertial SLAM with ORB-SLAM3 (Campos et al. 2021) using both default parameters and PSC estimated parameters. This result is tabulated in table 5.

The table compares the RMSE of ATE across several sequences using ORB-SLAM3 in both monocular visual inertial odometry (VIO) and monocular VO configurations. Each setup is evaluated using two parameter settings: the default parameters (as provided by EuRoC-MaV dataset (Burri et al. 2016)) and a PSC set estimated using the PSC pipeline. Across the board, bold values highlight superior performance in terms of lower RMSE.

Analyzing the monocular VIO, PSC parameters consistently outperform the default in nearly

all sequences, achieving lower RMSE values with all "*Machine Hall*" sequences except for "*MH04*". Similar trend is observed with the "*Vicon Room*" sequences except for "*V102*" and "*V101*". These observations show significant improvements in RMSE with PSC estimated parameters. Notably, "*MH02*" drops from 0.209 m to 0.065 m, and "*MH03*" from 0.193 m to 0.087 m. These observations demonstrate that PSC parameters generally enhance monocular VIO robustness and accuracy, even with more challenging sequences, which include aggressive motion.

Analyzing monocular-only VO shows much higher RMSE values overall compared to VIO, underscoring the importance of inertial fusion in trajectory estimation. However, PSC estimated parameters once again improve performance in most cases. The most dramatic improvements are seen in "*MH02*" (from 45.03 m to 10.55 m) and "*MH05*" (from 12.21 m to 3.33 m), showcasing the effectiveness of PSC in mitigating scale drift and motion ambiguity. For sequences "*MH01*", "*MH03*", "*MH04*", and "*V102*"–"*V202*", the PSC parameters also provide better or comparable accuracy. Exceptions include "*V101*", "*V201*", and "*V203*", where default parameters slightly outperform the PSC setup. However, the differences are often within the range of standard deviation, suggesting performance parity in these cases.

General trend that can be deciphered from Table 5 is that PSC parameters perform better in most "*Machine Hall*" sequences for both VIO and VO, but default IOPs performs better in more runs in the "*Vicon Room*" sequences. We believe this trend is because PSC IOPs is able to better leverage the feature-rich environment in the "*Machine Hall*" sequences to provide better feature extraction as compared to "*Vicon Room*" sequences which are mainly featureless.

3.5.2 Oxford Spires Dataset

Oxford-spires dataset (Tao et al. 2025) is a novel and diverse dataset which was captured using a handheld rig to test lidar odometry, and novel-view synthesis. Its diverse dataset includes sequences that are captured on a larger scale and indoors and outdoors. This unique transition between indoor and outdoor introduces variation in scene exposure which directly affects VO. The

Run No	Monocular-Inertial Odometry				Monocular Odometry			
	Default		PSC		Default		PSC	
	RMSE(m)	STD	RMSE(m)	STD	RMSE(m)	STD	RMSE(m)	STD
MH01	0.096	0.021	0.050	0.018	3.529	0.015	3.473	0.011
MH02	0.209	0.014	0.065	0.021	45.028	6.747	10.548	2.266
MH03	0.193	0.221	0.087	0.012	1.791	0.758	1.128	0.809
MH04	0.138	0.019	0.204	0.050	3.566	0.196	3.447	0.479
MH05	0.095	0.019	0.092	0.032	12.210	9.130	3.333	3.770
V101	0.033	0.004	0.032	0.001	0.959	0.008	0.968	0.013
V102	0.059	0.003	0.070	0.003	4.404	0.314	1.130	0.085
V103	0.070	0.006	0.049	0.017	0.855	0.029	0.447	0.060
V201	0.059	0.006	0.079	0.007	1.587	0.017	1.589	0.043
V202	0.063	0.004	0.047	0.019	0.885	0.427	0.665	0.328
V203	0.079	0.012	0.074	0.017	1.208	0.214	1.240	0.288

Table 5: Comparison of RMSE ATE running Visual-inertial SLAM using ORB-SLAM3 (Campos et al. 2021) with default parameters provided by (Burri et al. 2016) and the BA parameters that is determined by running PSC on the same data

ground truth (GT) poses are provided by laser scan matching and are in world frame T_{base}^{world} , hence this had to be transformed into camera from given by $T_{base}^{imu} \times T_{imu}^{cam0}$.

The results from running monocular visual SLAM are given in Table 6, which reports the ATE in terms of RMSE and standard deviation for monocular visual SLAM (ORB-SLAM3 (Campos et al. 2021)) on several sequences of the Oxford-spires dataset, comparing the default CF calibration parameters provided by the dataset (Tao et al. 2025) with PSC-refined parameters obtained.

For a majority of sequences, PSC calibration yields lower RMSE than the CF parameters. For example, "Keble College 02" and "Keble College 04" show reductions from 0.858 m to 0.418 m and 0.398 m to 0.273 m, respectively — improvements of nearly 50%. Similarly, in both "Observatory Quarter" sequences, the error is more than halved, dropping from 0.341 m to 0.137 m (for sequence "01") and 0.361 m to 0.159 m (for sequence "02"). Alongside lower RMSE, the corresponding standard deviations are also markedly reduced, indicating that BA calibration not only improves accuracy but also precision across camera poses.

The trend diverges for the "Blenheim Palace 01" and "Blenheim Palace 02" sequences. Here, CF parameters achieve lower RMSE (0.893 m and 2.249 m) compared to BA parameters (1.633 m and 2.814 m, respectively). We believe this is due to the blurry imagery in the sequence when the exposure time automatically increases to compensate for the low light when the handheld platform is captured inside the palace. This goes on to show that BA based IOPs and distortion coefficient leverage is diminished when the imagery has motion blur.

4 Conclusion

In this paper, we examined two widely used camera calibration approaches: the CF method, which relies on planar calibration targets and provides a guaranteed analytical solution, and PSC, which leverages 3D calibration targets, initial EOPs, and BA to jointly refine the IOPs, EOPs, and target coordinates. We outlined the model selection process and applied both methods to calibrate our cameras before evaluating their impact on downstream tasks including SfM, VO, and novel view synthesis.

		Default	PSC
Kemble College 02	RMSE(m)	0.858	0.418
	Std.(m)	1.218	0.211
Kemble College 04	RMSE(m)	0.398	0.273
	Std.(m)	0.204	0.234
Kemble College 05	RMSE(m)	1.531	1.520
	Std.(m)	0.176	0.405
Observatory Quarter 01	RMSE(m)	0.341	0.137
	Std.(m)(m)	0.158	0.016
Observatory Quarter 02	RMSE(m)	0.361	0.159
	Std.(m)(m)	0.187	0.017
Blenheim Palace 01	RMSE(m)	0.893	1.633
	Std.(m)(m)	0.566	0.625
Blenheim Palace 02	RMSE(m)	2.249	2.814
	Std.(m)	0.367	0.334

Table 6: Inverted comparison of RMSE ATE running Monocular Visual SLAM using ORB-SLAM3 (Campos et al. 2021) with default parameters provided by (Tao et al. 2025) and the BA parameters determined by running PSC on the same data.

Across all experiments, we showed that calibration parameters derived from PSC consistently outperformed those obtained from the CF method. In SfM reconstruction, PSC reduced mean absolute error from 0.209 m to 0.038 m, which is an order of magnitude improvement—while in VO experiments, trajectory drift dropped from over 16 m with CF calibration to nearly 1 m using PSC. Similarly, NERF reconstructions trained with glspc intrinsics achieved 5–7 dB higher PSNR, 0.2–0.3 higher SSIM, and up to 3–4× lower LPIPS, producing sharper, precise and more stable renderings.

Furthermore, we demonstrated that PSC can also be effectively applied to planar target datasets, such as those used in popular benchmarks, and still enhance the performance of downstream tasks — as shown in our VO experiments. Our findings highlight that CF calibration often fails to completely eliminate systematic errors

introduced by lens distortion, which is implicitly reflected in its estimated IOPs and distortion coefficients. While CF calibration has traditionally been adopted for non-critical computer vision tasks, our results strongly advocate for the broader use of PSC in safety- and precision-critical applications, such as autonomous navigation, mapping, and benchmark dataset preparation, where geometric accuracy is paramount.

Declarations

Funding: This study did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of interest: The authors declare no conflict of interest.

Ethics approval and consent to participate: Not applicable.

Data availability: This work uses both data collected with our mapping platform and publicly available datasets. The data captured by us is openly available at [this repository](#).

Materials availability: Not applicable.

Code availability: The bundle adjustment code is proprietary and cannot be shared. However, the camera parameters estimated in this work are provided for use with the released datasets.

References

- Brown DC (1971) Close-range camera calibration. *Photogrammetric Engineering* 37:855–866
- Burri M, Nikolic J, Gohl P, et al (2016) The euroc micro aerial vehicle datasets. *The International Journal of Robotics Research* <https://doi.org/10.1177/0278364915620033>, URL <http://ijr.sagepub.com/content/early/2016/01/21/0278364915620033>
- Campos C, Elvira R, Rodríguez JJG, et al (2021) Orb-slam3: An accurate open-source library for visual, visual-inertial,

- and multimap slam. *IEEE Transactions on Robotics* 37(6):1874–1890. <https://doi.org/10.1109/TRO.2021.3075644>
- Fraser C (2013) Automatic camera calibration in close range photogrammetry. *Photogrammetric Engineering & Remote Sensing* 79:381–388. <https://doi.org/10.14358/PERS.79.4.381>
- Furgale P, Rehder J, Siegwart R (2013) Unified temporal and spatial calibration for multi-sensor systems. In: 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp 1280–1286, <https://doi.org/10.1109/IROS.2013.6696514>
- Geiger A, Lenz P, Stiller C, et al (2013) Vision meets robotics: The kitti dataset. *International Journal of Robotics Research* 32:1231–1237. <https://doi.org/10.1177/0278364913491297>
- Grupp M (2017) evo: Python package for the evaluation of odometry and slam. <https://github.com/MichaelGrupp/evo>
- Kannala J, Brandt S (2006) A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 28(8):1335–1340. <https://doi.org/10.1109/TPAMI.2006.153>
- Kedilioglu O, Bocco TM, Landesberger M, et al (2021) Aruco: Enhanced aruco marker. In: 2021 21st International Conference on Control, Automation and Systems (ICCAS), pp 878–881, <https://doi.org/10.23919/ICCAS52745.2021.9650050>
- Lichti DD, Jarron D, Tredoux W, et al (2020) Geometric modelling and calibration of a spherical camera imaging system. *The Photogrammetric Record* 35:123–142. <https://doi.org/10.1111/phor.12315>, URL <https://onlinelibrary.wiley.com/doi/10.1111/phor.12315>
- Lowe DG (2004) Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision* 60(2):91–110. <https://doi.org/10.1023/B:VISI.0000029664.99615.94>, URL <https://doi.org/10.1023/B:VISI.0000029664.99615.94>
- Luhmann T, Robson S, Kyle S, et al (2014) Close-Range Photogrammetry and 3D Imaging. De Gruyter, Berlin, Boston, <https://doi.org/10.1515/9783110302783>, URL <https://doi.org/10.1515/9783110302783>
- Luhmann T, Fraser C, Maas HG (2016) Sensor modelling and camera calibration for close-range photogrammetry. *ISPRS Journal of Photogrammetry and Remote Sensing* 115:37–46. <https://doi.org/https://doi.org/10.1016/j.isprsjprs.2015.10.006>, URL <https://www.sciencedirect.com/science/article/pii/S0924271615002361>
- Maybank SJ, Faugeras OD (1992) A theory of self-calibration of a moving camera. *International Journal of Computer Vision* 8(2):123–151
- Olson E (2011) Apriltag: A robust and flexible visual fiducial system. In: 2011 IEEE International Conference on Robotics and Automation, pp 3400–3407, <https://doi.org/10.1109/ICRA.2011.5979561>
- Pan L, Barath D, Pollefeys M, et al (2024) Global Structure-from-Motion Revisited. In: European Conference on Computer Vision (ECCV)
- Ramezani M, Wang Y, Camurri M, et al (2020) The newer college dataset: Hand-held lidar, inertial and vision with ground truth. In: 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp 4353–4360, <https://doi.org/10.1109/IROS45743.2020.9340849>
- Rehder J, Nikolic J, Schneider T, et al (2016) Extending kalibr: Calibrating the extrinsics of multiple imus and of individual axes. In: 2016 IEEE International Conference on Robotics and Automation (ICRA), pp 4304–4311, <https://doi.org/10.1109/ICRA.2016.7487628>
- Richardson A, Strom J, Olson E (2013) Aprilcal: Assisted and repeatable camera calibration. In: 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp 1814–1821, <https://doi.org/10.1109/IROS.2013.6696595>

- Schönberger JL, Frahm JM (2016) Structure-from-motion revisited. In: Conference on Computer Vision and Pattern Recognition (CVPR)
- Schönberger JL, Price T, Sattler T, et al (2016a) A vote-and-verify strategy for fast spatial verification in image retrieval. In: Asian Conference on Computer Vision (ACCV)
- Schönberger JL, Zheng E, Pollefeys M, et al (2016b) Pixelwise view selection for unstructured multi-view stereo. In: European Conference on Computer Vision (ECCV)
- Schubert D, Goll T, Demmel N, et al (2018) The tum vi benchmark for evaluating visual-inertial odometry. In: 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, pp 1680–1687, <https://doi.org/10.1109/IROS.2018.8593419>, URL <https://ieeexplore.ieee.org/document/8593419/>
- Sun P, Kretzschmar H, Dotiwalla X, et al (2020) Scalability in perception for autonomous driving: Waymo open dataset. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition pp 2443–2451. <https://doi.org/10.1109/CVPR42600.2020.00252>
- Tancik M, Weber E, Ng E, et al (2023) Nerfstudio: A modular framework for neural radiance field development. In: ACM SIGGRAPH 2023 Conference Proceedings, SIGGRAPH '23
- Tao Y, Muñoz-Bañón MÁ, Zhang L, et al (2025) The oxford spires dataset: Benchmarking large-scale lidar-visual localisation, reconstruction and radiance field methods. International Journal of Robotics Research
- Tarrio JJ, Pedre S (2015) Realtime edge-based visual odometry for a monocular camera. In: Proceedings of the IEEE International Conference on Computer Vision (ICCV)
- The MathWorks I (2025) Camera Calibration Toolbox. Natick, Massachusetts, United States, URL <https://www.mathworks.com/help/symbolic/>
- Triggs B (1998) Autocalibration from planar scenes. In: Burkhardt H, Neumann B (eds) Proceedings of the 5th European Conference on Computer Vision, Part I. Springer-Verlag, Freiburg, Germany, pp 89–105, <https://doi.org/10.1007/BFb0055661>
- Vo AV, Laefer DF, Byrne J (2021) Optimizing urban lidar flight path planning using a genetic algorithm and a dual parallel computing framework. Remote Sensing 13(21). <https://doi.org/10.3390/rs13214437>, URL <https://www.mdpi.com/2072-4292/13/21/4437>
- Wang W, Hu Y, Scherer SA (2020) Tartanvo: A generalizable learning-based vo. In: Conference on Robot Learning, URL <https://api.semanticscholar.org/CorpusID:226226415>
- Xu W, Cai Y, He D, et al (2022) Fast-lio2: Fast direct lidar-inertial odometry. IEEE Transactions on Robotics 38(4):2053–2073. <https://doi.org/10.1109/TRO.2022.3141876>
- Zhang Z (2000) A flexible new technique for camera calibration. IEEE Transactions on Pattern Analysis and Machine Intelligence 22:1330–1334. <https://doi.org/10.1109/34.888718>, URL <http://ieeexplore.ieee.org/document/888718/>
- Zheng C, Xu W, Zou Z, et al (2025) Fast-lio2: Fast, direct lidar-inertial-visual odometry. IEEE Transactions on Robotics 41:326–346. <https://doi.org/10.1109/TRO.2024.3502198>