Evaluating Smartphone Camera Calibration Configurations and their Effects on SLAM

Rhea Joyce Zambra[a], Saurav Uprety [a], Ananya Vishwanath [a], Favour Iheanyichukwu[a], Hongzhou Yang[a]

[a]Dept. of Geomatics Engineering, University of Calgary, Alberta, Canada – honyang@ucalgary.ca

Keywords: Smartphone Camera Calibration, Calibration Grid Type, Calibration Grid Size, Distortion Model

ABSTRACT:

Smartphones have become increasingly viable for photogrammetric and simultaneous localization and mapping (SLAM) applications due to their portability and widespread availability. However, the repeatability of smartphone camera calibration remains a concern, as intrinsic orientation parameters (IOPs) can vary significantly between calibration attempts due to innate software and hardware corrective mechanisms. This study investigates the impact of calibration grid type, grid size, and distortion modelling on smartphone camera calibration uncertainty and its downstream effects on positioning accuracy in monocular SLAM. Using three smartphone models (iPhone 14 and two Google Pixel 7 devices), we conducted a comprehensive analysis of 24 calibration configurations processed through Kalibr. The estimated IOPs were then applied in a monocular ORB-SLAM3 pipeline, and the resulting trajectories were compared against a high-precision integrated LiDAR inertial ground truth. These findings provide insights into optimizing smartphone calibration setups, which has effects on SLAM-based applications in mobile mapping, robotics, and augmented reality.

1. INTRODUCTION

Due to their ubiquity and compact size, smartphones have increasingly been researched for their potential in accurate photogrammetric (Patonis, 2024) and simultaneous localization and mapping applications (SLAM) (Zambra *et al.*, 2024). In traditional photogrammetric surveys, camera calibration is performed in a prior and separate process to determine the interior orientation parameters (IOPs), which are stable in metric and semi-metric cameras due to their specialized design. However, smartphones often exhibit variability in IOPs across different calibration attempts (Jasińska *et al.*, 2023). These variabilities can introduce uncertainties in the downstream processes of positioning and 3D reconstruction, as techniques like SLAM and structure from motion (SfM) often assume constant IOPs between frames.

Camera calibration determines the IOPs – focal length, principal point, and distortion coefficients – by analyzing a series of images to establish an accurate relationship between a 3D scene and its 2D projection onto an image. Contemporary calibration methods typically optimize an overdetermined system of image points with known 2D-to-3D correspondence, such as with a known set of 3D points through self-calibration, a geometrically known planar grid, or geometric constraints on natural features through SfM. Current state-of-the-art methods follow Zhang's flexible approach of using multiple views of a planar grid to estimate the camera intrinsic and extrinsic parameters (Zhang, 2000). This paradigm is adopted in the popular libraries of Kalibr (Furgale *et al.*, 2012), OpenCV, and the MATLAB Camera Calibration Toolbox due to the setup simplicity. Additionally, target-based calibration outperforms targetless methods, such as iKalibr (Chen *et al.*, 2024) and ATiPE, (Barazzetti *et al.*, 2012) in terms of repeatability due to the high contrast and easily detectable nature of the target grids. However, despite the extensive bodies of work on target-based calibration, the impact of grid size, grid type, and distortion model on the repeatability of smartphone camera calibration has received limited attention in literature. Compared to traditional high-grade photogrammetric cameras, smartphones are subject to greater IOP variability, due to factors including lens tolerances, autofocus mechanisms, and software-based corrections. Understanding these inconsistencies is important for improving smartphone-based SLAM applications**,** particularly in robotics, augmented reality (AR), and mobile mapping.

To address this gap, a comprehensive analysis of smartphone camera calibration is performed, evaluating the effects of grid size, grid type, and distortion models on measurement uncertainty and SLAM positioning. Experiments will be conducted using three smartphones (iPhone 10, two Google Pixel 7 devices) to account for variations across different operating systems and same-model inconsistencies, as well as to identify consistent calibration trends across devices.

2. RELATED WORK

Prior research has explored various aspects of camera calibration, but the influence of these factors on the repeatability of IOP estimates and the resulting positioning accuracy for smartphones has not been extensively studied. The following subsections review relevant literature on key factors affecting calibration performance. By analyzing these factors, this study seeks to provide insights into their role in enhancing calibration consistency and improving positioning accuracy in smartphone SLAM applications.

2.1 Calibration Grid Type

The choice of the type of calibration grid has shown to have effects on IOP estimation. In 2007, Mallon and Whelan conducted a comparative study of two full-visibility calibration grids – the checkerboard and circular dot grid – and concluded that the latter is more susceptible to distortion bias, degrading the predicted pixel position of the features (Mallon and Whelan, 2007). However, full-visibility grids require that the entire target remain within the camera's field of view, limiting the range of calibration poses and reducing coverage around the image edges. To address this, partially visible grids, such as Aprilgrid (Olson, 2011) and ChArUco, have been created with uniquely identifiable markers within the pattern. These grids allow calibration even when only a portion of the grid is visible and provide improved estimation of lens distortion, particularly at the periphery of the image.

Given the variability in smartphone camera characteristics, it is important to investigate how different grid types influence calibration repeatability and downstream SLAM performance. This study extends previous work by evaluating both a full-visibility (checkerboard) and a partial-visibility grid (Aprilgrid) to assess their impact on positioning accuracy.

2.2 Calibration Grid Size

Grid size is another factor affecting camera calibration, particularly in smartphone cameras with autofocus. Autofocus mechanisms dynamically adjust the lens position, leading to subtle variations in focal length across different calibration attempts. Smaller grids necessitate a shorter focusing distance compared to larger grids, introducing inconsistencies in focal length estimation. These focal length variations, in turn, can propagate errors into the estimation of extrinsic parameters (EOPs), thereby biasing the estimated IOPs (Ricolfe-Viala and Esparza, 2021).

Ricolfe-Viala and Esparza's study on autofocus cameras demonstrated that even when bundle adjustment yields reprojection errors close to zero, biases in intrinsic and extrinsic parameter estimates can persist. This highlights the limitations of standard calibration pipelines when applied to autofocus systems. Building on their findings, this study investigates how calibration grid size impacts IOP estimate uncertainty, repeatability, and positioning accuracy in smartphone-based SLAM by using two significantly different grid sizes. By quantifying the propagation of calibration errors into the positioning domain, we aim to provide insights into optimizing grid size selection for more reliable smartphone SLAM.

2.3 Camera Distortion Modelling

Another common source of smartphone camera calibration error is incorrect lens distortion modeling, which has adverse effects on 3D reconstruction accuracy (Patonis, 2024). The most widely used model in photogrammetry is the Brown-Conrady model (Duane, 1971), which assumes that radial and tangential distortion are functions of the distance from the optical center. Kannala and Brandt introduced an alternative approach through an equidistant distortion model designed for wide-angle lenses, which models distortion as a function of angular deviation from the optical center (Kannala and Brandt, 2006). While this approach has been widely applied in fisheye and panoramic cameras, its effectiveness for smartphones remains underexplored. Moreover, most studies on distortion modelling focus on image reprojection accuracy and 3D reconstruction, rather than its effects on positioning accuracy in SLAM applications. This study aims to address this gap by evaluating two different distortion models (radial-tangential and equidistant) and analyzing their influence on smartphone-based monocular trajectory estimation.

2.4 Experimental Setup

Building on the insights from previous studies, this research implements an experimental setup to systematically evaluate the impact of calibration grid type, grid size, and distortion model on IOP estimate repeatability and positioning accuracy. The experimental configurations, summarized in Table 1, follow a standardized calibration pipeline using Kalibr, generating eight sets of IOP estimates per smartphone, resulting in a total of 24 configurations across three devices.

Each set of IOP estimates will be analyzed for uncertainty based on the final covariance estimate, including focal length, principal point, and distortion coefficients. The estimated IOPs will then be applied as fixed parameters in monocular SLAM through ORB-SLAM3, chosen for its robustness in monocular SLAM. The resulting camera trajectories will be compared against a LiDAR-inertial based ground truth. The positioning accuracy will be assessed using the root mean square error (RMSE) of the absolute positioning error (APE), to quantify the deviations between the estimated and reference trajectories.

Table 1. Configurations for smartphone camera calibration.

Grid Size \ Grid Type	Aprilgrid	Checkerboard
Small	pinhole rad-tan	pinhole rad-tan
	pinhole equi-dist	pinhole equi-dist
Large	pinhole rad-tan	pinhole rad-tan
	pinhole equi-dist	pinhole equi-dist

3. PRELIMINARIES

3.1 Kalibr Camera Intrinsic Estimation

In Kalibr, the Levenberg-Marquardt (LM) algorithm is used to optimize the IOPs and EOPs of the camera by minimizing the reprojection error e, which measures the sum of the difference between the observed feature points p_i and predicted location $\hat{p}_i$:

$$e = \sum_i \left|\left|p_i - \hat{p}_i\right|\right|^2 \quad (1)$$

This is done by using a projection model π, which maps the i^{th} 3D point P_i to image coordinates $p_i = (x, y)$ using the camera intrinsic matrix K and its EOP θ:

$$\hat{p}_i = \pi(K, \theta, P_i) \quad (2)$$

From the observations p_i (automatically extracted 2D image points) and the predicted locations $\hat{p}_i$, the residuals r_i are calculated as the difference:

$$r_i = p_i - \hat{p}_i \quad (3)$$

To update the camera parameters iteratively, the Jacobian J of the residuals with respect to the parameters are computed:

$$J = \frac{\partial r_i}{\partial \theta} \quad (4)$$

The Levenberg-Marquardt algorithm then uses J to adjust the camera parameters θ in each iteration, minimizing e. Once optimization converges, the uncertainty in the estimated camera parameters is obtained from the covariance matrix C_θ and variance of residuals σ^2:

$$C_\theta = \sigma^2 (J^T J)^{-1} \quad (5)$$

where the diagonal elements of C_θ provide the uncertainty (standard deviation) for the camera EOPs and IOPs.

3.2 Lens Distortion Models

3.2.1 Radial-Tangential Model: The Brown-Conrady model is widely used in photogrammetry and assumes that distortion is a function of the radial distance from the optical center. It models both radial distortion (causes straight lines to curve) and tangential distortion (decentering effects due to misalignment):

$$x_{distorted} = x(1 + k_1 r^2 + k_2 r^4 + \cdots) \\ y_{distorted} = y(1 + k_1 r^2 + k_2 r^4 + \cdots) \quad (6)$$

$$x_{distorted} = x + [2p_1 xy + p_2(r^2 + 2x^2)] \\ y_{distorted} = y + [p_1(r^2 + 2y^2) + 2p_2 xy] \quad (7)$$

where k_1, k_2 are the radial distortion coefficients, p_1, p_2 are the tangential distortion coefficients, and $r = \sqrt{x^2 + y^2}$ is the radial distance from the principal point.

3.2.2 Equidistant Model: On the other hand, the Kannala-Brandt equidistant model assumes distortion as being proportional to the angle θ from the optical axis:

$$\theta = arctan\left(\frac{r}{f}\right) \quad (8)$$

$$\delta r = k_1\theta + k_2\theta^3 + k_3\theta^5 + k_4\theta^7 + \cdots \quad (9)$$

where k_1, k_2, k_3, k_4 are equidistant distortion coefficients, r is the radial distance, f is the focal length, and δr is the radial distortion from the lens.

Although higher order coefficient parameters can be calculated for both models, Kalibr outputs $[k_1\ k_2\ r_1\ r_2]$ for radial-tangential and $[k_1\ k_2\ k_3\ k_4]$ for equidistant distortion calibration.

4. METHODOLOGY

4.1 Smartphone Camera Calibration

Images were captured by hand using smartphones in front of four distinct calibration grids (Table 2): small Aprilgrid (sAG), large Aprilgrid (bAG), small checkerboard (sCH), and large checkerboard (bCH) (Figure 1). A standardized calibration routine was followed for all calibration data collection, using the SensorLogger app (Tszheichoi, 2024) with the frame rate set to 30 frames per second. Using three smartphones — an iPhone 14 (iP14), Google Pixel 7 #1 (GP7P1), and Google Pixel 7 #4 (GP7P4) — this process resulted in twelve distinct camera calibration datasets. Each dataset was processed through Kalibr for camera calibration, performed twice for two distortion models: radial-tangential (PR) and equidistant (PE), yielding a total of twenty-four distinct camera intrinsic calibration configurations.

4.2 Smartphone and LiDAR Trajectory

To assess positioning accuracy, smartphone image data and the UNIX timestamp (through SensorLogger) was collected concurrently with ground truth data. The ground truth positions were acquired using a Velodyne 16-channel LiDAR and an EPSON EG320N inertial measurement unit (IMU) from a NovAtel SPAN (Figure 2), with the data processed via tight-coupling with FAST-LIO2 (Xu *et al.*, 2021) (Figure 3). For the smartphone trajectories, ORB-SLAM3 (Campos *et al.*, 2021) was chosen as the SLAM processor due to its robustness in visual-SLAM (Figure 4). All sensors and smartphones were mounted on a cart, which was used to carry the equipment during data collection. Data from all three smartphones, along with the ground truth, were captured during the same run in an indoor, texture-rich laboratory environment. The ground truth trajectory for the selected data segment is shown in Figure 5.

Table 2. Configuration of each calibration grid.

	Aprilgrid		Checkerboard	
	Small	**Big**	**Small**	**Big**
Columns	6	6	4	8
Rows	6	6	6	9
Target Size (cm)	2.80	9.00	3.60	10.00

Figure 1. The four calibration grids used: big Aprilgrid (top left), small Aprilgrid (top right), big checkerboard (bottom left), and small checkerboard (bottom right).

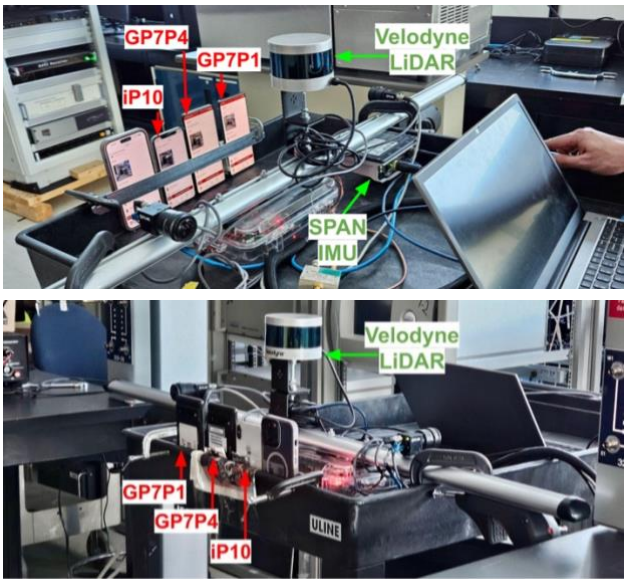

Figure 2. Data collection platform used for indoor positioning, front (top) and back (bottom) view.

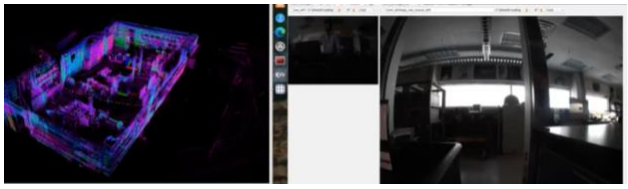

Figure 3. FAST-LIO2 processing of LiDAR-inertial data.

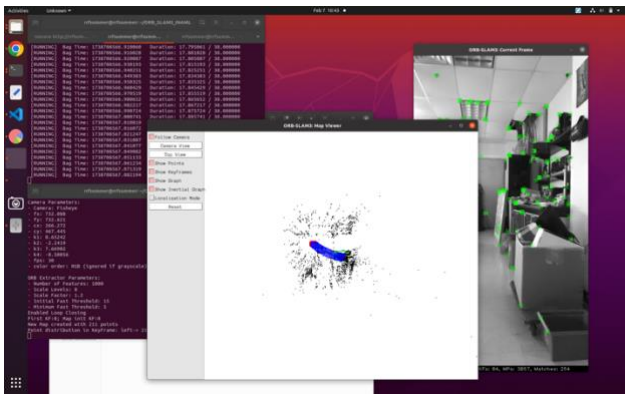

Figure 4. ORB-SLAM3 processing of smartphone data.

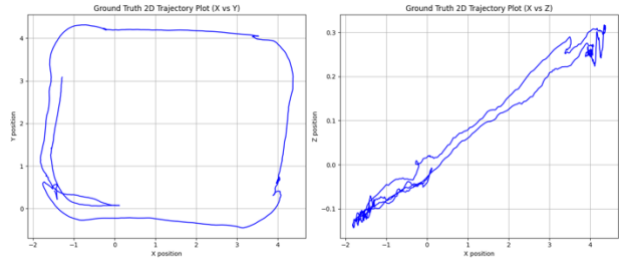

Figure 5. LiDAR-IMU ground truth trajectory via FAST-LIO2.

5. RESULTS & DISCUSSION

5.1 Smartphone Camera Calibration Uncertainty

The uncertainty in the camera IOP estimates is quantified through the standard deviations provided in the output from the covariance C_θ. Combinations of the standard deviations σ_{avg} are calculated using the pooled standard deviation equation:

$$\sigma_{avg} = \sqrt{\frac{\sigma_1^2 + \sigma_2^2 + \cdots + \sigma_n^2}{n}} \quad (10)$$

where σ_n^2 is the n^{th} variance, and n is the sample size.

The results indicate that the following configurations reduce uncertainty: larger grid size, Aprilgrid type, equidistant distortion modelling for the principal point, and radial-tangential distortion modelling for focal length (Tables 3-5). These findings are likely due to several factors: the larger calibration grids enhance robustness against motion blur, which improves feature detection; the Aprilgrid's partial visibility provides more data for calibration and allow for more parallax between views; the equidistant distortion model, which lacks a non-symmetric component, simplifies the estimation of the principal point; and the radial-tangential distortion model has a more complex corrective model than equidistant, leading to more accurate focal length estimates.

Table 3. Comparison of IOP estimate uncertainty for the two grid types across all calibration configurations.

	Aprilgrid	Checkerboard
σ_{f_x} [px]	1.19	1.72
σ_{f_y} [px]	1.20	1.71
σ_{c_x} [px]	0.91	1.56
σ_{c_y} [px]	0.96	1.71

Table 4. Comparison of IOP estimate uncertainty for the two distortion models across all calibration configurations.

	Equidistant	Radial-Tangential
σ_{f_x} [px]	1.56	1.39
σ_{f_y} [px]	1.56	1.39
σ_{c_x} [px]	0.95	1.54
σ_{c_y} [px]	0.89	1.74

Table 5. Comparison of IOP estimate uncertainty for the two grid sizes across all calibration configurations.

	Large	Small
σ_{f_x} [px]	1.05	1.80
σ_{f_y} [px]	1.07	1.80
σ_{c_x} [px]	0.89	1.58
σ_{c_y} [px]	0.90	1.74

Additionally, Table 6 shows the combined uncertainty breakdown by smartphone, showing that GP7P1 had the least uncertainty, and that the same model smartphones (GP7P1 and GP7P4) had sub-pixel level difference in uncertainty.

Table 6. Comparison of IOP estimate uncertainty for the three smartphones across all calibration configurations.

	GP7P1	GP7P4	iP10
σ_{f_x} [px]	1.25	1.46	1.69
σ_{f_y} [px]	1.27	1.46	1.68
σ_{c_x} [px]	1.11	1.15	1.53
σ_{c_y} [px]	1.29	1.32	1.53

It can be noted that the grid size creates the largest differentiation in uncertainty, closely followed by the calibration grid type. However, it is worth noting that while large grids reduce uncertainty, it causes the residuals to be stretched along the phone's long axis after optimization (Figure 6). This suggests that calibration is more sensitive to distortions in this direction, possibly due to more exaggerated camera motion. The anisotropic distribution indicates that the model may not fully compensate for non-uniform distortions, thus larger grids may reduce uncertainty overall but could introduce localized errors along the longer axis.

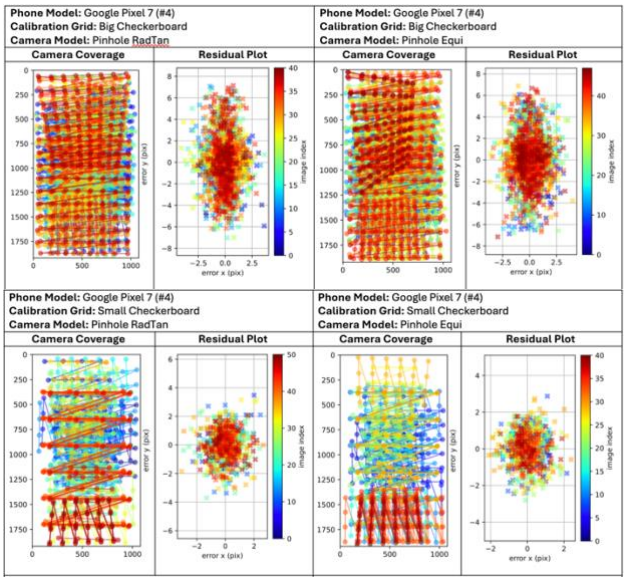

Figure 6: Anisotropic distribution of reprojection errors (top) after calibration using large grids compared to small grids.

As for distortion modeling, the uncertainty is greatly reduced when using an Aprilgrid instead of a checkerboard (Table 7), as its partial-visibility allowance provides better camera coverage at the edges of the images (Figure 7). Using a larger calibration grid also reduces uncertainty by improving robustness against motion blur and ensuring better coverage of distortions by increasing feature size (Table 8).

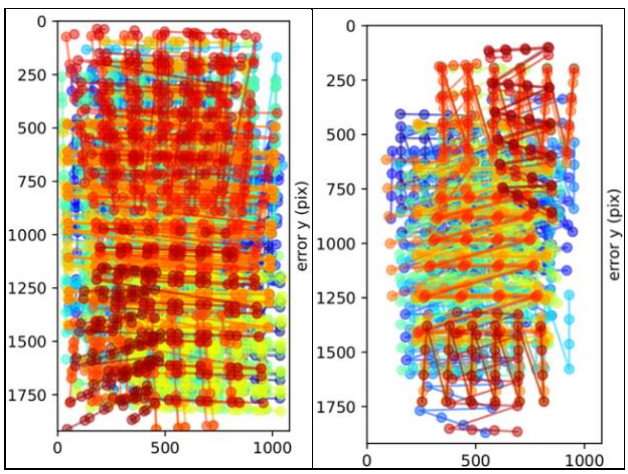

Figure 7: Camera coverage plot for the iPhone 10: small Aprilgrid (left), and small checkerboard (right).

Table 7. Comparison of distortion parameter uncertainty for the two grid types across 24 calibration configurations.

	Equidistant		Radial-Tangential	
	AG	CH	AG	CH
k1	0.00319	0.01603	0.00132	0.00273
k2	0.02809	0.15245	0.00224	0.00582
k3	0.09955	0.59804		
k4	0.12106	0.81406		
r1			0.0002	0.0005
r2			0.0002	0.0004

Table 8. Comparison of distortion parameter uncertainty for the two grid sizes across all calibration configurations.

	Equidistant		Radial-Tangential	
	Small	Big	Small	Big
k1	0.01546	0.00888	0.00272	0.00149
k2	0.14677	0.08020	0.00438	0.00261
k3	0.57797	0.28820		
k4	0.79136	0.35267		
r1			0.00047	0.00027
r2			0.00041	0.00025

Additionally, Table 9 shows the combined distortion uncertainty breakdown by smartphone, showing that GP7P4 had the least distortion uncertainty, and that the same model smartphones (GP7P1 and GP7P4) have similar levels of uncertainty.

Table 9. Comparison of distortion parameter uncertainty for the three smartphones across all calibration configurations.

	Equidistant		
	GP7P1	**GP7P4**	**iP10**
k1	0.01066	0.01091	0.01562
k2	0.09025	0.08933	0.16074
k3	0.30464	0.29309	0.66855
k4	0.34814	0.32670	0.16074
	Radial-Tangential		
	GP7P1	**GP7P4**	**iP10**
k1	0.00171	0.00162	0.00287
k2	0.00297	0.00265	0.00652
r1	0.00035	0.00032	0.00043
r2	0.00028	0.00026	0.00044

5.2 Smartphone Camera Intrinsics

The variability between the IOP estimates show that IOP variations due to different calibration setups can exceed **50 pixels** for focal length (f_x, f_y) and **35 pixels** for the principal point (c_x, c_y) (Table 10). This highlights the importance of calibrating for individual phones and not for individual smartphone models. Additionally, comparisons between the two same-model smartphones (GP7P1 vs. GP7P4) suggest that using a large Aprilgrid and equidistant distortion modelling may help maintain IOP stability between calibration sets (Table 11). Finally, the standard deviation analysis for calibration grid type and size indicates that larger calibration grids yield more consistent results (Table 12), whereas differences between checkerboard and Aprilgrid has inconclusive results (Table 13).

Table 10. Peak-to-peak variation in IOP estimates for each smartphone.

		f_x [px]	f_y [px]	c_x [px]	c_y [px]
iP10	Min	1464.18	1465.24	531.53	911.25
	Max	1498.34	1496.92	548.65	943.10
	Diff.	**34.17**	**31.68**	**17.12**	**31.85**
GP7P1	Min	1328.55	1331.60	518.33	943.39
	Max	1382.44	1378.33	541.44	978.87
	Diff.	**53.89**	**46.73**	**23.11**	**35.48**
GP7P4	Min	1311.75	1311.15	537.07	927.66
	Max	1359.27	1357.03	559.57	946.49
	Diff.	**47.53**	**45.87**	**22.50**	**18.83**

Table 11. Summary of same-model absolute differences (GP7P1-GP7P4) for each IOP.

Calibration Configuration	f_x [px]	f_y [px]	c_x [px]	c_y [px]
PE_sAG	55.98	54.15	29.16	21.21
PE_bAG	2.10	1.22	3.77	3.86
PE_sCH	22.96	20.86	27.45	50.71
PE_bCH	19.07	16.10	6.53	2.52
PR_sAG	16.22	16.07	10.55	10.74
PR_bAG	7.81	8.02	14.42	16.31
PR_sCH	36.87	34.23	14.07	32.19
PR_bCH	14.66	13.44	2.20	9.25

Table 12. Summary of standard deviation of IOP estimates by calibration grid size across all calibration configurations.

	Small	Big
$STDEV_{f_x}$ [px]	15.94	13.92
$STDEV_{f_y}$ [px]	15.21	12.69
$STDEV_{c_x}$ [px]	7.53	6.05
$STDEV_{c_y}$ [px]	11.38	6.32

Table 13. Summary of standard deviation of IOP estimates by calibration grid type across all calibration configurations.

	AG	CH
$STDEV_{f_x}$ [px]	14.58	15.34
$STDEV_{f_y}$ [px]	14.22	13.79
$STDEV_{c_x}$ [px]	6.87	6.79
$STDEV_{c_y}$ [px]	7.37	10.73

Additionally, the standard deviation of the estimated IOPs for each smartphone is summarized in Table 14, showing that GP7P1 had the least variability across calibration tests, while iP10 had the most homogeneous standard deviations for the x and y directions.

Table 14. Summary of standard deviation of IOP estimates by smartphone across all calibration configurations.

	GP7P1	GP7P4	iP10
$STDEV_{f_x}$ [px]	12.10	19.00	14.01
$STDEV_{f_y}$ [px]	12.09	16.52	14.06
$STDEV_{c_x}$ [px]	5.49	7.80	7.45
$STDEV_{c_y}$ [px]	9.13	11.18	7.61

5.3 Smartphone Camera Positioning Evaluation

To assess positioning accuracy, the RMSE of the APE is used as the primary metric. APE quantifies the deviation between estimated camera poses $T_{est,i}$ and the ground truth poses $T_{gt,i}$ at corresponding timestamps. Before computing APE, an Umeyama Sim(3) alignment is applied to correct for scale differences in the monocular SLAM trajectory, which is inherently arbitrary in scale. Example aligned trajectories for each smartphone is displayed in Figure 8
. Mathematically, the APE for a given pose i is defined as:

$$\mathrm{APE_i} = \left|\left|\mathrm{T_{est,i}} - \mathrm{T_{gt,i}}\right|\right| \quad (11)$$

and the RMSE of the APE is:

$$APE_{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} APE_i^2} \quad (12)$$

which represents the average deviation between the ground truth and estimated pose across all timestamps N.

The average RMSE is calculated for each calibration setup group, and results indicate that the small grid size, checkerboard grid type, and equidistant distortion modelling lead to lower positioning errors (Tables 15-17). The grid configurations are particularly beneficial because they allow for greater depth variation during the calibration process. Firstly, smaller grids enable the user to position the camera very close to the grid, which is useful for close-range 3D reconstruction. Secondly, the checkerboard grid type, unlike the Aprilgrid, remains recognizable at larger distances, making it effective for a more varied calibration process. Additionally, equidistant distortion is a symmetrical, less complex distortion model that may be more fitting to the smartphone's wide field of view. Together, these factors contribute to more accurate positioning estimates, as they provide better calibration coverage across various depths and perspectives.

Table 15. A comparison of the APE RMSE value between the two calibration grid sizes.

	Small	Big
APE RMSE [cm]	38.8	41.5

Table 16. A comparison of the APE RMSE value between the two calibration grid types.

	Aprilgrid	Checkerboard
APE RMSE [cm]	41.6	38.7

Table 17. A comparison of the APE RMSE value between the two camera distortion models.

	Equidistant	Radial-Tangential
APE RMSE [cm]	35.1	45.2

Additionally, the APE RMSE values by smartphones are summarized in Table 18, showing that the iPhone 10 had the least positioning errors, closely followed by GP7P1.

Table 18. A comparison of the APE RMSE value between the three smartphones.

	GP7P1	GP7P4	iP10
APE RMSE [cm]	38.7	43.4	38.3

6. CONCLUSION AND OUTLOOK

This study examined how calibration grid type, grid size, and camera distortion model affect smartphone camera calibration uncertainty and the effects on monocular SLAM positioning. The results suggest that no single calibration setup is universally optimal; instead, parameter selection should be tailored to the application, with a focus on repeatability and depth variation.

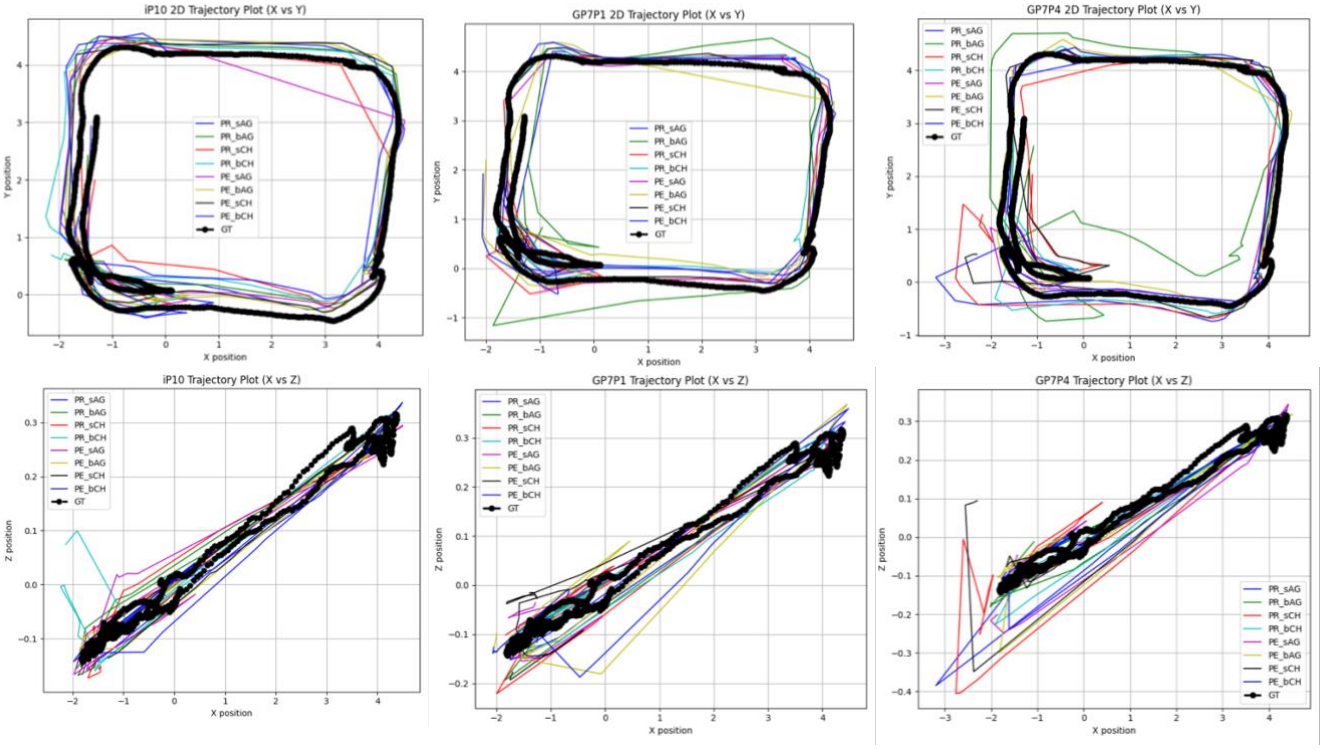

Figure 8. The aligned trajectory plots across all calibration configurations for the iPhone 10 (left), Google Pixel 7 #1 (middle) and Google Pixel 7 #4 (right). This shows how each IOP set from different calibration configurations causes differences in the trajectory estimate for each phone.

Grid Size: Larger grids improve robustness against motion blur, making them preferable for applications requiring consistent intrinsic and distortion modeling, reducing uncertainty by approximately 0.7 pixels for focal length and principal point. However, they are more prone to axis-specific errors, especially in dynamic environments. In contrast, smaller grids are better suited for close-range depth estimation in indoor SLAM, reducing positioning errors by 2.7 cm. However, they are more susceptible to motion blur, leading to increased uncertainty.
Grid Type: The Aprilgrid outperforms the checkerboard grid in modeling intrinsic parameters, reducing uncertainty by approximately 0.6 pixels for focal length and principal point. It also provides better distortion modeling due to its partial visibility and increased image coverage, making it advantageous for cameras with complex distortion. However, the Aprilgrid is less effective in scenarios requiring a wide depth range, such as SLAM in diverse environments. The checkerboard grid, though more limited in visibility, is better suited for applications requiring depth variation, reducing positioning errors by 2.9 cm.
Distortion Model: The equidistant distortion model proves more effective for smartphone SLAM, likely due to its reduced dependence on radial distance and simpler model, lowering positioning errors by 10.1 cm. While the radial-tangential model is well-suited for general camera lenses, the equidistant model should be consided more for smartphone SLAM applications.

For future work, several key areas warrant further investigation:
1. Extended smartphone testing: Future research should evaluate a wider range of smartphone models, considering differences across manufacturers, camera hardware generations, and repeated calibrations on the same device to assess consistency.
2. Temporal analysis of IOP stability: Smartphone camera calibration repeatability may degrade over time due to environmental factors, component aging, and software updates
3. Integration of IMU: Monocular SLAM lacks scale, limiting metric accuracy. Investigating the impact of calibration on IMU extrinsics and metric positioning could expand the use of smartphones in high-precision SLAM applications.

Addressing these areas will further refine smartphone-based photogrammetry and SLAM methodologies, enhancing the reliability of mobile mapping, robotics, and augmented reality applications.

Acknowledgements

We gratefully acknowledge the support of the Alberta Innovates in this project.

References

Barazzetti, L., Mussio, L., Remondino, F., Scaioni, M., 2012. TARGETLESS CAMERA CALIBRATION. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. XXXVIII-5-W16.

Campos, C., Elvira, R., Rodríguez, J.J.G., M. Montiel, J.M., D. Tardós, J., 2021. ORB-SLAM3: An Accurate Open-Source Library for Visual, Visual–Inertial, and Multimap SLAM. IEEE Trans. Robot. 37, 1874–1890.

Chen, S., Li, X., Li, S., Zhou, Y., Yang, X., 2024. iKalibr: Unified Targetless Spatiotemporal Calibration for Resilient Integrated Inertial Systems.

Duane, C.B., 1971. Close-range camera calibration. Photogramm. Eng, 37(8), pp.855-866

Furgale, P., Barfoot, T.D., Sibley, G., 2012. Continuous-time batch estimation using temporal basis functions. IEEE International Conference on Robotics and Automation, pp. 2088–2095.

Jasińska, A., Pyka, K., Pastucha, E., Midtiby, H.S., 2023. A Simple Way to Reduce 3D Model Deformation in Smartphone Photogrammetry. Sensors 23, 728.

Kannala, J., Brandt, S.S., 2006. A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses. IEEE Trans. Pattern Anal. Mach. Intell. 28, 1335–1340.

Mallon, J., Whelan, P.F., 2007. Which pattern? Biasing aspects of planar calibration patterns and detection methods. Pattern Recognit. Lett. 28, 921–930.

Olson, E., 2011. AprilTag: A robust and flexible visual fiducial system. IEEE International Conference on Robotics and Automation, pp. 3400–3407.

Patonis, P., 2024. Comparative Evaluation of the Performance of a Mobile Device Camera and a Full-Frame Mirrorless Camera in Close-Range Photogrammetry Applications. Sensors 24, 4925.

Ricolfe-Viala, C., Esparza, A., 2021. The Influence of Autofocus Lenses in the Camera Calibration Process. IEEE Trans. Instrum. Meas. 70, 1–15.

Tszheichoi, 2024. SENSOR LOGGER, Version 1.31.3. Kelvin Choi. https://www.tszheichoi.com/sensorlogger (1 April 2024).

Xu, W., Cai, Y., He, D., Lin, J., Zhang, F., 2021. FAST-LIO2: Fast Direct LiDAR-inertial Odometry.

Zambra, R.J., Uprety, S., Lee, R., Yang, H., 2024. Smartphone HD Map Updates Using Monocular-Inertial ORB-SLAM3 and Gaussian Splatting. Proceedings of the 37th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2024), pp. 1885–1900.

Zhang, Z., 2000. A flexible new technique for camera calibration. IEEE Trans. Pattern Anal. Mach. Intell. 22, 1330–1334.