Accuracy estimation of 3D model creation using ToF cameras integrated into smartphones

Ivan A. Romanenko, Alexandr G. Chibunichev

Moscow State University of Geodesy and Cartography, Russia - romanenko.inc@gmail.com, agchib@mail.ru

Keywords: Time-of-Flight (ToF) camera, smartphone, camera calibration, spatial test object, camera internal orientation elements, 3D modeling accuracy.

Abstract:

This work investigates the accuracy of determining object point coordinates from images acquired with a Time-of-Flight (ToF) camera integrated into an iPhone 15 Pro smartphone. Calibration of such a camera consist of two parts: direction calibration and distance calibration. For classical direction calibration (determining the interior orientation parameters), we propose a spatial test object consisting of a plane-parallel plate with known dimensions that can move perpendicularly to its surface. A universal photogrammetric device (SPR—Romanovski stereo projector) was used for this. The plate movements were recorded using the meter with an accuracy of 0.1mm. At the same time, the camera is stationary and a series of images is taken with uniform elements of external orientation. Therefore, they can be considered a single image. This is how a spatial test object is formed. For distance calibration, we propose determine the scale factor of the model built from pixel-coordinate measurements and corresponding distances. Our results demonstrate that classical ToF camera calibration, which corrects point directions by accounting for lens distortion, does not improve point-coordinate accuracy in ToF-derived models. Model accuracy significantly improves when applying the scale factor determined from control points or known distances on the object.

1.Introduction

Time-of-Flight (ToF) cameras are rapidly being adopted across industries, enabling new opportunities for 3D model generation. They are used in industrial quality control, medical 3D reconstruction of organs, AR/VR object tracking, and in mobile devices for enhanced portrait photography, virtual object overlay, space scanning (He & Chen, 2019; Qiu et al., 2023; Liu et al., 2018; Nakagawa & Taguchi, 2020), and facial recognition (Stotko et al., 2019; Weinmann et al., 2020) (Figure 1). However, the potential of smartphone-integrated ToF cameras for 3D modeling tasks remains underexplored, particularly regarding their accuracy compared to professional systems such as the Microsoft Kinect (Kurillo et al., 2022).

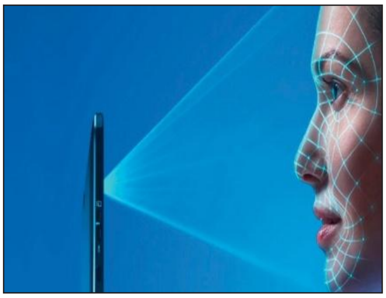

Fig. 1 Example of ToF camera application in a smartphone for face recognition.

Time-of-Flight (ToF) cameras are devices that measure the distance to objects using the delay time of a light pulse. ToF cameras use two measurement methods:

1) Direct measurement (dToF): a short IR pulse is emitted towards an object, the sensor records the time of its return. The distance D is calculated by the equation:

$$D = \frac{c \times \Delta t}{2}, \quad (1)$$

where c is the speed of light, Δt is the delay time.

2) Indirect measurement (iToF): modulated IR light (sinusoid) is reflected from the object, the phase shift Δφ between the sent and received signal allows to calculate the distance:

$$D = \frac{c \times \Delta\varphi}{4\pi f}, \quad (2)$$

where f is the modulation frequency.

Each pixel of the sensor measures the distance to the object forming a depth map - a matrix of values where each number corresponds to the distance at a particular point in the scene. The resolution of the map depends on the sensor used in the device (Table 1).

Product Name	Resolution	Type
Microsoft Azure Kinect	1024x1024	iToF
Intel Real Sense LiDAR L515	1024x768	dToF
Sony DepthSense IMX556	640x480	iToF
iPhone TrueDepth camera	640x360	iToF

Table 1: Typical representatives of ToF cameras

The advantages of ToF cameras include high measurement speed (real-time mode), compactness, operation in low-light conditions. There are also limitations, namely the fusion of external light, sunlight reduces accuracy, not all materials reflect IR pulses (transparent, mirror or black surfaces), short range (most models do not work within the range above 10 meters, and also sensitive to temperature conditions (noise may appear).

Current methods for improving ToF camera accuracy include hybrid approaches such as FloatingFusion (Meuleman et al.,

Publisher's note: Copernicus Publications has not received any payments from Russian or Belarusian institutions for this paper.

This contribution has been peer-reviewed.

2023), which aims to improve the accuracy of depth estimation using time-of-flight (ToF) cameras and stabilized stereo cameras in smartphones. Modern smartphones are equipped with multimodal systems incorporating ToF sensors and multiple RGB cameras, enabling the creation of depth maps for computer vision and photogrammetry tasks. However, accurate depth mapping is difficult due to the low resolution of ToF sensors and optical image stabilization (OIS), which changes the position of the main camera lens, disturbing the geometric relationships between the sensors.

To address this problem, the authors propose a FloatingFusion approach that combines data from a ToF sensor and a wide-angle RGB camera. The key feature is an online calibration that allows automatic detection of external, internal parameters and distortions of the stabilized main camera based on dense 2D/3D correspondences. This enables fusion processing of the data through a correlation volume combining information from both the stereo pair and the ToF sensor. Real scenes processed with NeRF (Neural Radiance Fields) with ToF data control are used to train the model.

Experiments were conducted on a proprietary dataset obtained using high accuracy Kinect Azure and on test scenes with 200 images. The results show that the proposed method provides better accuracy compared to existing approaches such as TöRF, NerfingMVS, CVD and others. FloatingFusion is particularly good at recovering fine details and edges of objects, showing robustness against noise and artifacts of ToF sensors. It has also been shown that ignoring OIS leads to a significant reduction in reconstruction quality, emphasizing the importance of online calibration.

Thus, FloatingFusion opens up new possibilities for creating highly accurate 3D models using smartphone cameras, especially in the face of limited sensor power and challenging surveying conditions.

In contrast to this approach, this paper focuses on classical ToF camera calibration using a spatial test object and a scale factor. The results show that lens distortion correction does not improve model accuracy, while the introduction of a scale factor reduces the RMS significantly. This is consistent with the findings of Meuleman et al. (2023) on the importance of calibration, but emphasizes the specificity of smartphone embedded sensors, where factory calibration may limit the effectiveness of additional corrections.

In this study, we evaluate the accuracy of 3D modeling using the iPhone 15 Pro's ToF camera by comparing coordinates of a test object's points with those obtained via ToF measurements.

2.Methodology

The operating principle of a ToF camera is based on measuring the time-of-flight of light to and from the object. The phase difference between the emitted and received signals is recorded by a photonic mixing device (PMD) sensor. The camera illuminates the object with near-infrared LEDs. The distance $\mathbf{D_M}$ (Figure 2) to each object point is computed for each pixel at coordinates x,y on the PMD matrix using the phase difference. The object point coordinates M(X, Y, Z) are then computed from pixel coordinates m(x, y) and $\mathbf{D_M}$:

$$X_M = D_M \cos\nu \sin\varphi ; \quad Y_M = D_M \sin\nu ;$$

$$Z_M = D_M \cos\nu \cos\varphi . \tag{3}$$

where the horizontal angle φ and vertical angle ν are derived from pixel coordinates ***(xy)*** of point m in the ToF camera image:

$$tg\varphi = \frac{x}{f} ; \qquad tg\nu = \frac{y}{f}\cos\varphi . \tag{4}$$

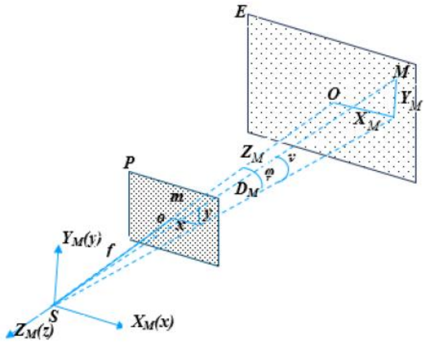

Fig.2 The principle of calculating the coordinates of object points based on the results of measurements of ToF camera image point coordinates and distances.

Thus, for each pixel with known PMD coordinates and corresponding distance, we reconstruct a 3D point cloud in the object coordinate system SXYZ. Model accuracy depends on interior orientation parameters (focal length, principal point coordinates, and lens distortion) and distance-measurement precision. Interior orientation parameters can be obtained via classical photogrammetric calibration, while distance calibration has been addressed by Lindner & Kolb (2006) and Sobers et al. (2011) through comparison with known object distances. Since distance measurement errors affect the model scale, we propose determining the model scale factor t using multiple control points of the spatial test object.

As the test object, a metallic plane-parallel rectangular plate (267 × 120 mm) was mounted on an analog photogrammetric device (SPR) capable of translating along the Z-axis with submillimeter precision. Corner coordinates were measured with a steel ruler. The iPhone 15 Pro was clamped approximately 35 cm from the plate.

The plate was translated in 10 mm increments (0.1 mm precision) and photographed in two configurations: parallel (Figure 3) and inclined (Figure 5) relative to the camera. Each series of fixed-camera images with constant exterior parameters constitutes a single multi-view dataset, simulating a spatial test object of dimensions 267×120×100 mm with 40 control points (Figure 4).

https://doi.org/10.5194/isprs-archives-XLVIII-2-W9-2025-227-2025 | © Author(s) 2025. CC BY 4.0 License.

Fig.3 Capturing the plate with a fixed smartphone camera

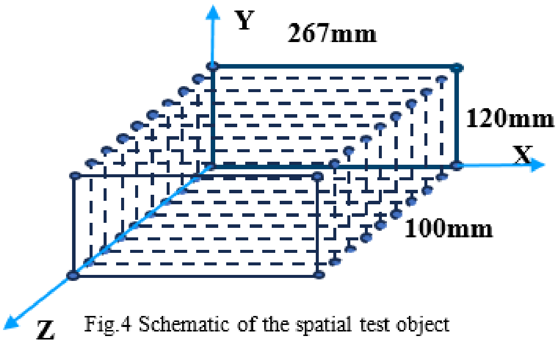

Fig.4 Schematic of the spatial test object

Fig.5 Capturing the plate with a fixed smartphone camera tilted in relation to the plate

All images taken from the same camera position share the same exterior orientation elements, so they can be considered as a single image that depicts a spatial test object.

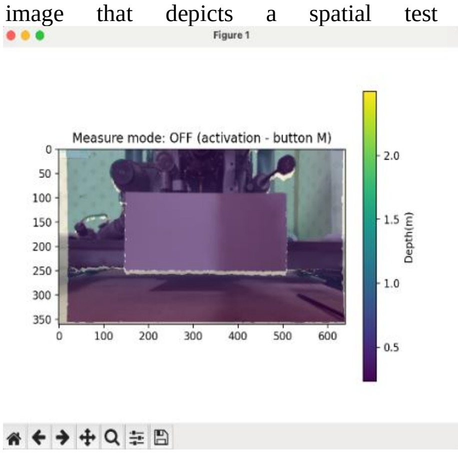

Fig.6 Interface of the program for measuring pixel coordinates of ToF camera image points and corresponding distances

A python program was developed to perform measurements. The program requires an image obtained from the ToF camera in the '.exr' format and an RGB camera image obtained at the same moment with the ToF camera in the '.jpg' format. The program combines the images for convenient measurements. Coordinate values on the image are displayed in pixels, depth value - in meters. The measurement mode is activated by the M key, after which clicking on the image the pixel coordinates x, y and distance d in the specified pixel in meters are written to the program memory. Measurement results are unloaded to a text file for further processing. (Figure 6) .

Next, the camera was calibrated using measurements of the coordinates of the plate points in all 20 images, counting them as two single images (Govorov, A.V. et al., 2020). Calibration was performed using the Photomod software.

Standard distortion correction formulas were applied:

$$\begin{rcases} d_x = (x-x_0)(r^2k_1+r^4k_2+r^6k_3)+(r^2+2(x-x_0)^2)p_1+2(x-x_0)(y-y_0)p_2; \\ d_y = (y-y_0)(r^2k_1+r^4k_2+r^6k_3)+2(x-x_0)(y-y_0)p_1+(r^2+2(y-y_0)^2)p_2, \end{rcases} \quad (5)$$

In order to facilitate the calculation of the corrected coordinates, a program was developed using the Python programming language using equations (3)–(4) –(5) (Fig.6). The program receives as input a text file containing the point name, initial pixel coordinates and distance. The operator is required to enter the camera calibration parameters, following which the calculation of XM, YM and ZM model coordinates becomes available. Furthermore, a window for calculating the distance between two points was developed, with the option of comparing with the reference value, if available, for approximate estimation of the obtained coordinates.

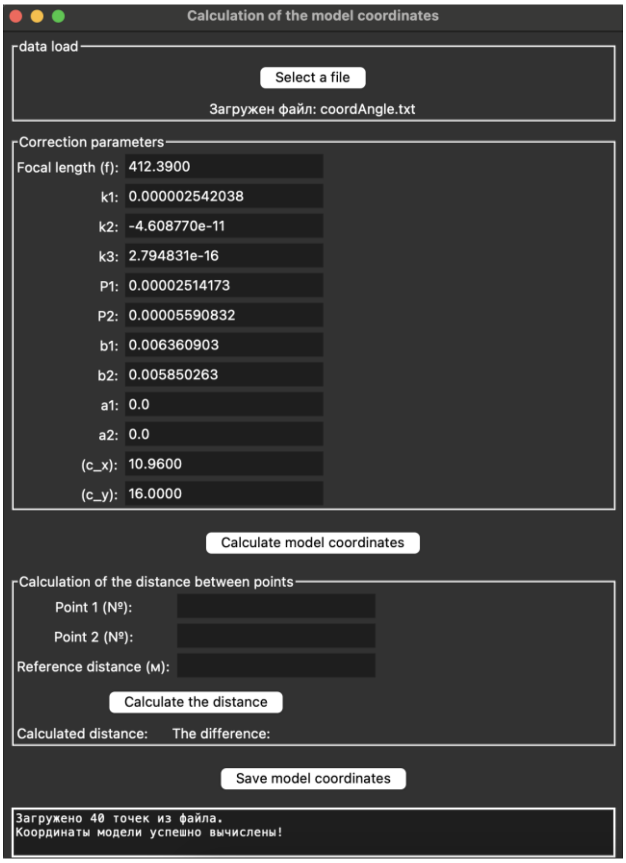

Figure 6. Model coordinate computation program

Corrected coordinates were used to reconstruct the spatial model via equations (3), and absolute orientation applied using control points:

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} X_0 \\ Y_0 \\ Z_0 \end{pmatrix} + \mathbf{A}_M \begin{pmatrix} X_M \\ Y_M \\ Z_M \end{pmatrix} t \qquad (6)$$

Here X, Y, Z — coordinates of the object point in the object coordinate system (Fig.4); X_M, Y_M, Z_M — coordinates of the corresponding model point in the model coordinate system (3); X_0, Y_0, Z_0 — coordinates of the beginning of the model coordinate system in the object coordinate system; A_M matrix of rotation of the model coordinate system relative to the object coordinate system; t - scale factor of the model.

Below are the results of camera calibration and estimation of the accuracy of building the model of the test object.

3.Results

Camera calibration was performed simultaneously using two images (parallel to the plate and oblique). A total of 80 reference points (40 points on one image and 40 on the second image). As a result of the camera calibration, the following values (Table 2) of the interior orientation elements were obtained.

Parameter values are presented in pixels

f	x_0	y_0	k_1	k_2	k_3	p_1	p_2
Radial distortion							
400	2.5	-4.9	1.8E-06	-1.3E-11	-	-	-
Full calibration							
412.3	10.9	16	2.5E-06	-4.6E-11	2.8E-16	2.5E-05	5.6E-05

Table 2 Camera calibration results

The accuracy of the object model construction was assessed by:

1) differences in the distances measured on the plate and calculated from the coordinates of the corresponding points of the model. A total of 40 distances along the X-axis, 40 along the Y-axis, and 8 along the Z-axis were used (Fig. 4).

2) differences between coordinates of reference points after absolute model orientation and coordinates of corresponding points in the object coordinate system (Fig.4). A total of 40 points were used.

A program was written in the Python programming language to calculate the RMS from the differences of distances on the plate along the axes (Fig.7). The program for calculations requires a text file that contains the point name, X, Y and Z coordinate. Also additionally the user must enter the value of the scale factor, its value will be multiplied with each coordinate in the original data. If the scaling factor value is equal to 1, the original coordinate values will be processed. After calculations the results can be downloaded in '.xslx' format..

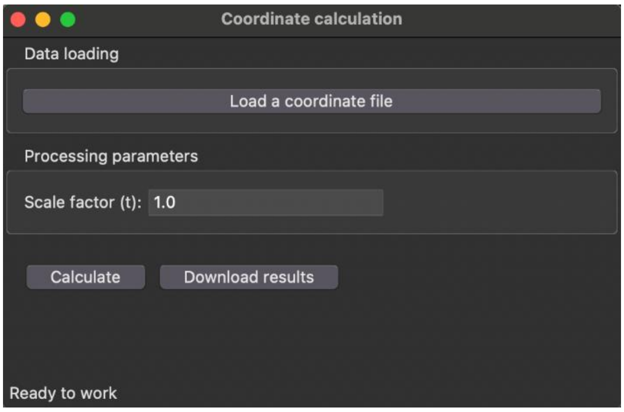

Figure 7. RMS computation program

The corresponding root mean square errors (RMS) are given in Table 3 and are provided in meters. t is the scale factor calculated during the absolute orientation of the model.

	RMS on the divergence of the plate distances along the axes: (m)			RMS by coordinates of reference points (40 pcs.) (m)			t
	X	Y	Z	m_X	m_Y	m_Z	
From the image parallel to the plate.							
Uncalibrated	0.017	0.008	0.003	0.004	0.002	0.004	1.055
Uncalibrated + t	0.007	0.003	0.005				
Radial distortion	0.037	0.017	0.002	0.004	0.002	0.006	1.129
Radial distortion + t	0.009	0.004	0.007				
Full calibration	0.031	0.014	0.003	0.004	0.002	0.005	1.119
Full calibration + t	0.008	0.003	0.007				
From the image oblique to the plate							
Uncalibrated	0.004	0.003	0.005	0.002	0.002	0.003	0.991
Radial distortion	0.019	0.008	0.005	0.002	0.003	0.004	1.072
Radial distortion + t	0.003	0.004	0.005				
Full calibration	0.014	0.006	0.005	0.002	0.002	0.004	1.049
Full calibration + t	0.003	0.004	0.006				

Table 3 Estimation of model building accuracy by reference points

Analyzing the results of estimation of model building accuracy given in Table 3 we can draw the following conclusions. The main source of errors of object model building are errors of distance measurement by the ToF camera, which are largely taken into account when introducing the scale factor t. Calibration of the camera at different variants of lens distortion gives approximately the same results. This indicates that this camera is most likely factory calibrated and the output is distortion-corrected images. Therefore, when using this camera to create 3D models, it is recommended to determine the scale factor of the model. Besides, the coordinates of the test object model points obtained from a oblique image are somewhat more accurate than the corresponding coordinates obtained from a parallel image. Therefore, it is recommended to take photos such objects at different angles with respect to the object.

4.Conclusion

Classical ToF camera calibration, which corrects directions to object points by taking into account lens distortion, does not result in improved accuracy of model point coordinates obtained from the ToF camera image. The accuracy of the object model is significantly improved by taking into account the scale factor of the model, which can be determined from reference points or known distances on the object.

The obtained results of estimation of the accuracy of object point coordinates determination using ToF camera of iPhone 15 Pro should be considered preliminary. More extensive studies with different cameras should be performed.

In the future, it is possible to integrate ToF camera with RGB-sensor by determining their mutual orientation in order to obtain textured 3D-models. In addition, the possibility of operating the ToF camera in low light or total darkness should be investigated, which expands its application in the field.

References

He, Y., & Chen, S. (2019). Recent advances in 3D data acquisition and processing by time-of-flight camera. IEEE Access, 7, 12495–12510.

Kurillo, G., Hemingway, E., Cheng, M.-L., & Cheng, L. (2022). Evaluating the accuracy of the Azure Kinect and Kinect v2. Sensors, 22(7), 2469.

Liu, Y., Gao, W., & Hu, Z. (2018). Geometrically stable tracking for depth images based 3D reconstruction on mobile devices. ISPRS Journal of Photogrammetry and Remote Sensing, 143, 222–232.

Nakagawa, M., & Taguchi, M. (2020). Moving object classification using multilayer laser scanning with space subdivision framework. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, V-4-2020, 131–138.

Qiu, Z., Martínez-Sánchez, J., & Arias, P. (2023). Evaluation and comparison of different time of flight cameras for outdoor applications. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, X-1/W1-2023, 17–23.

Sobers, L. X. Francis, S. G. Anavatti, & M. Garratt. (2011). Reconstructing the geometry of an object using 3D TOF camera. Proceedings of MFCIST, April 2011, https://doi.org/10.1109/MFCIST.2011.5949516.

Stotko, P., Weinmann, M., & Klein, R. (2019). Albedo estimation for real-time 3D reconstruction using RGB-D and IR data. ISPRS Journal of Photogrammetry and Remote Sensing, 150, 213–225.

Weinmann, M., Jäger, M. A., Wursthorn, S., Jutzi, B., Weinmann, M., & Hübner, P. (2020). 3D indoor mapping with the Microsoft HoloLens: Qualitative and quantitative evaluation by means of geometric features. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, V-1-2020, 173–180.

Govorov, A. V., Chibunichev, A. G., & Makarov, S. B. (2020). Study of digital camera calibration on a flat test object. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLIII-B2-2020, 29–33. https://doi.org/10.5194/isprs-archives-XLIII-B2-2020-29-2020

Lindner, M., & Kolb, A. (2006). Lateral and depth calibration of pmd-distance sensors, in Advances in Visual Computing. Springer, pp. 524–533.

Meuleman, A., Kim, H., Tompkin, J., & Kim, M. H. (2022). FloatingFusion: Depth from ToF and Image-stabilized Stereo Cameras . arXiv preprint arXiv:2210.02785. https://arxiv.org/abs/2210.02785