Stereo Calibration of Heterogeneous RGB-ToF Camera for Robotics Applications

Sagar Dhatrak
Einfochips (An Arrow Company)
Pune, India
sagar.dhatrak@einfochips.com

Sunil Kumar*
Einfochips (An Arrow Company)
Noida, India
sunil.kumar5@einfochips.com

Aayush Shah
Einfochips (An Arrow Company)
Ahmedabad, India
aayush.shah1@einfochips.com

Amit Gupta
Einfochips (An Arrow Company)
Noida, India
amit.gupta@einfochips.com

2025 6th International Conference on Artificial Intelligence, Robotics and Control (AIRC) | 979-8-3315-4348-8/25/$31.00 ©2025 IEEE | DOI: 10.1109/AIRC64931.2025.11077476

***Abstract*—RGB-Depth (or RGB-D) cameras are increasingly being used in robotic and vision systems, from service robotics to industrial robotics applications. To effectively use RGB-D cameras in such applications, accurate intrinsic and extrinsic calibration parameters are essential. These parameters must be precise to achieve reliable and accurate results. Stereo calibration of two identical cameras is relatively straightforward compared to heterogeneous cameras (e.g., ToF-RGB) and has been studied by various researchers. This research article presents a methodology for the stereo calibration of two heterogeneous cameras with differing image resolutions and pixel sizes. It proposes a method to make pixel sizes and resolutions of two cameras identical to ensure accurate calibration. The infrared (IR) image captured by the ToF camera often suffers from low brightness and contrast, which can degrade the quality of captured images. As a result, this can lead to inaccurate calibration parameters during the calibration process. To address this, the IR images are enhanced to improve the accuracy of calibration parameters. Apart from this, this paper outlines the proper calibration process, with the effectiveness of the proposed approach assessed in terms of reprojection error. Furthermore, the calibrated images have been tested using MATLAB and the OpenCV Library to analyze the reprojection error.**

Keyword—RGB-D camera, calibration, vision systems, reprojection error

I. Introduction and Related work

In recent years, RGB-D cameras have attracted great attention due to their low cost and ability to capture color and depth information. The availability of low-cost depth sensors in combination with regular RGB cameras, often embedded in the same device called an RGB-D camera, has provided a complete and instantaneous representation of both the color information and the 3D structure of the surrounding environment. These sensors rely on Time-of-Flight (ToF) technology [1] or structured infrared light [2], to estimate depth, while color information is extracted using an RGB camera. RGB-D cameras have been recently utilized in an increasing number of applications such as SLAM and navigation [3], 3-D scene parsing [4], tracking [5], object recognition and localization [6], augmented reality [7], 3-D shape scanning [8], and object and gesture recognition [9]. All these applications require precise knowledge of intrinsic parameters (which specify the camera's internal characteristics, such as focal length and optical center) along with distortion coefficients, as well as extrinsic parameters, such as the relative position and orientation between the RGB and depth sensors. RGB-D devices are frequently factory calibrated to obtain the calibration parameters, which are stored in non-volatile memory. Unfortunately, the quality of this calibration is only suitable for gaming purposes [16]. For instance, with the default configuration, the acquired point clouds may exhibit inaccurate correlation between depth and RGB data due to a lack of precise alignment between the camera and the depth sensor. As a result, an accurate calibration approach is required for robust applications. There are still several challenges in calibrating RGB-D cameras, particularly in terms of precision and usability.

Many researchers have developed different techniques for RGB-D camera calibration. A calibration method based on severe distortion that requires only the camera to take one image in a specific location has been developed by Huang and Soung [10]. They performed an experiment on a plane target that has low equipment requirements. Zhou et al. [11] proposed a new method for improving RGB-D camera calibration, particularly for depth (infrared) cameras, by collecting calibration images using a checkerboard under an auxiliary infrared light source. They also employed the proposed depth correction model to directly calibrate the depth image's distortion. To estimate the relative poses between RGB-D cameras with a minimum overlapping FoV, Liu et al. [12] introduced a new approach that uses well-matched key points provided by a feature descriptor-based calibration pattern to calibrate the RGB-D cameras' extrinsic parameters. To address the issue of significant disparities in camera separation, Su et al. [13] presented an algorithm that employs a spherical object to test various transformation functions, such as rigid transformation, polynomial transformation, and manifold regression for modeling camera extrinsic calibration. To determine the relative poses between RGB and depth cameras without necessitating an overlapping field of view, Yus et al. [14] introduced a technique that matches and extracts scene lines from both RGB and depth cameras, then applies geometric constraints to determine the relative poses between the sensors. Chen et al. [15] introduced a reliable and precise technique for calibrating Kinect-like sensors using motion capture systems. The suggested technique estimates the intrinsic and extrinsic parameters of each camera and generates a collection of pixel-wise depth correction models using heteroscedastic Gaussian processes. Basso et al. [16] introduced a novel approach that provides both the intrinsic and extrinsic parameters, requiring the user to gather data in a minimally structured environment. To calibrate all the geometric properties of RGB-D sensors, Darwish et al. [17] proposed a two-step calibration process. To remove the systematic inaccuracy caused by the baseline between the IR camera and projector, the in-factory calibration settings were first changed. Secondly, the distortion and systematic errors caused by the infrared camera and projector were adjusted using a combined distortion model. A new camera calibration technique was presented by Semeniuta [18] that is based on the removal of outliers and the numerical analysis of the probability distributions of the calibration parameters. A unique and user-friendly calibration technique for RGB-D cameras was introduced by Staranowicz et al. [20]. It uses a sphere's image projection as the only calibration object. Unlike the current checkerboard-based calibration method, the suggested algorithm extracts visual information without requiring user participation.

Researchers have conducted numerous experiments on RGB-D camera calibration to demonstrate the advantages of the proposed algorithm. However, there are still certain limitations, such as intrinsic parameter variations, depth sensor noise, infrared (IR) brightness issues, geometric distortion, field of view discrepancies, and occlusions. Among these, the brightness issue in infrared (IR) images during calibration is the primary concern. Therefore, this research article presented a work that demonstrates the calibration of two heterogeneous cameras with differing image resolutions and pixel sizes. The IR image captured by the ToF camera often suffers from low brightness and contrast, which can degrade the quality of captured images. To address this, the IR images are enhanced to improve their quality. This work presents detailed steps and a straightforward approach for the calibration of two different cameras. The rest of this paper is organized as follows: Section II provides an overview of the experimental approach. Section III describes the methodology of the work. Section IV presents experimental results, and Section V concludes the paper with future recommendations.

II. Experimental Approach

A. Experimental Setup

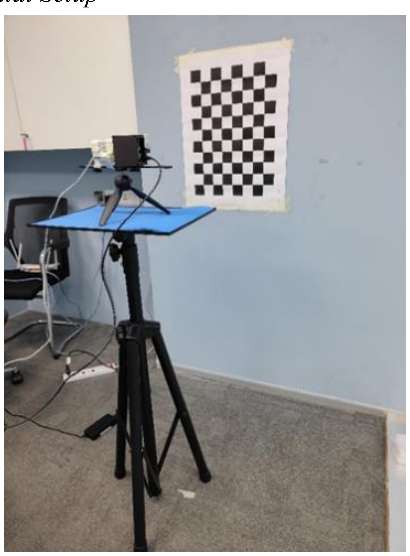


Fig. 1. Experimental Setup for stereo calibration of ToF-RGB camera

The experimental setup consists of an ADI ToF camera and a Lenovo camera, both mounted side by side on a tripod, as shown in Fig.1. The camera on the left-hand side is the Lenovo RGB camera, and the camera on the right-hand side is the ToF camera. Experiments were conducted in an indoor environment, free from any sunlight interference. This experimental setup, called the ToF-RGB camera, is designed for two tasks: the first is the single-camera calibration of the ADI ToF and Lenovo RGB cameras, and the second is to perform stereo camera calibration to estimate the geometric relationship between the two cameras.

B. Experimental Design

A scheme of experiment design is shown in Fig.2 [21]. The checkerboard pattern, with dimensions of 10x7 and each square with size 45mm, is placed on the wall as depicted in Fig. 2. Stereo camera setup is now facing towards the wall where a checkboard pattern is placed. This setup is positioned at different locations with varying orientations as shown in Fig. 2. The setup is positioned at different viewpoints, such as left center, right center, left, right, top look down, bottom lock up, etc. at a distance of 0.87m from checkerboard pattern to capture IR image and RGB image pairs. These images were subsequently captured at distances of 0.92m and 0.97m from the checkerboard pattern.

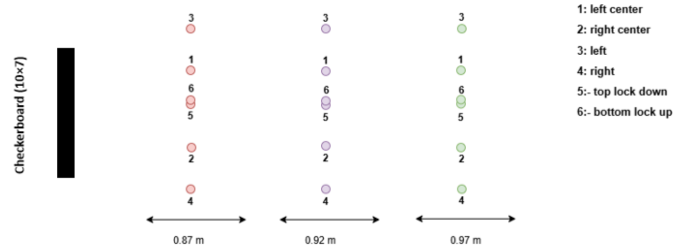


Fig. 2. Capturing IR and RGB images from multiple viewpoints for calibration

III. Methdology

Camera calibration is essential for accurately estimating the intrinsic parameters of the camera, removing lens distortions (such as radial and tangential), and determining extrinsic parameters, i.e, the position of the camera relative to the scene or other cameras. The calibration process was carried out using MATLAB and OpenCV to estimate intrinsic and extrinsic parameters for both RGB, ToF and ToF-RGB cameras.

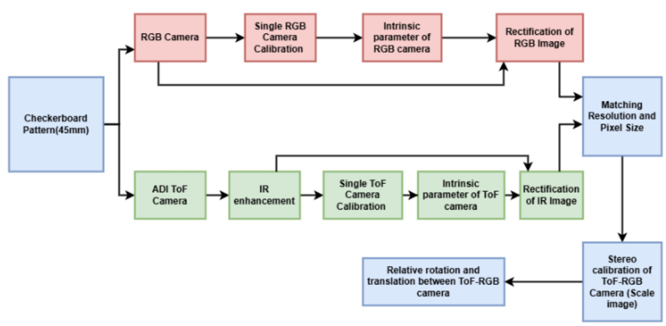


Fig. 3. An overview of the calibration approach.

The calibration methodology consists of the following steps: dataset collection, IR image enhancement, single camera calibration, and stereo camera calibration. Fig. 3 depicts an overview of the proposed calibration approach. Image datasets were captured for both the ToF camera and RGB camera. Infrared (IR) images are captured by the ToF camera, and RGB images are captured by the RGB camera. For each camera, multiple images of a checkerboard pattern were captured with different viewpoints and under uniform lighting conditions. Pairs of IR and RGB images were captured for stereo calibration of the ToF-RGB camera setup.

In IR image enhancement, the IR images captured by ToF camera have low brightness and contrast; hence it will affect calibration accuracy. Before doing calibration related to the ToF camera, these IR images need to be enhanced to improve calibration accuracy. These IR images are subsequently enhanced through normalization and transformed into 16-bit grayscale images [20]. It first determines the minimum and maximum values from the input image and normalizes the values based on the range between them. Each pixel's value is then mapped to a grayscale value between the minimum color value and the maximum value, scaling according to the normalized difference between the pixel and the minimum IR value. This conversion results in a 16-bit grayscale image suitable for further processing or display.

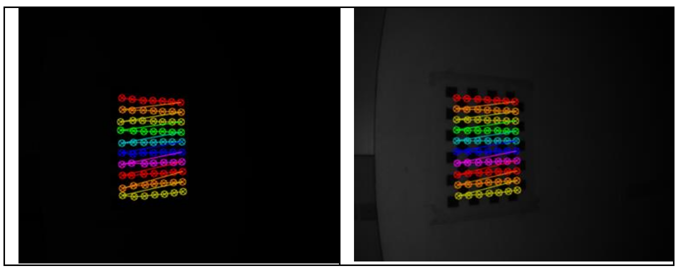


Fig. 4. The left image shows corners detected in the original, unenhanced image, while the right image shows corners detected in the enhanced image.

Fig. 4 shows two IR images before and after image enhancement. In the left-hand side image, the corner of the checkerboard pattern does not detect properly. Whereas, on the other side (right side image), the checker pattern is detected properly after image enhancement. The calibration of the RGB camera and ToF camera are performed individually using both MATLAB and OpenCV libraries to estimate the intrinsic parameters and distortion coefficients with the help of checkerboard patterns. For ToF camera calibration, enhanced IR images are used. This step provides the geometric parameters of individual camera such as focal length, principal point, and distortion coefficients. The stereo calibration of the ToF-RGB camera was performed to extract the extrinsic parameters, i.e rotation and translation between the ToF and RGB cameras using both MATLAB and OpenCV libraries using checkerboard patterns. This method uses the dataset collected for stereo calibration, but instead of the original IR images, it uses enhanced IR images. After stereo calibration, the performance of the calibration was evaluated by analyzing the reprojection error for each camera and the ToF-RGB stereo camera setup. A comparative analysis was then conducted between MATLAB and OpenCV. Results of the calibration process were visualized to check the quality of image alignment, and enhanced IR image processing was analyzed for its impact on checkerboard corner detection and reprojection accuracy.

IV. Experimental Results

After collecting the datasets, the pairs of IR and RGB images were used for stereo calibration with OpenCV and MATLAB tools. First, individual camera calibration was carried out, followed by stereo camera calibration. The accuracy of the single and stereo camera calibration processes was evaluated using reprojection error analysis.

A. Calibration of RGB, ToF and ToF-RGB Camera and Alignment

The RGB camera calibration was carried out to estimate the intrinsic parameters of the RGB camera. The estimated calibration matrix is shown below.

$$K_{rgb} = \begin{bmatrix} fx_{rgb} & 0 & Cx_{rgb} \\ 0 & fy_{rgb} & Cy_{rgb} \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 707.56 & 0 & 637.98 \\ 0 & 707.62 & 343.46 \\ 0 & 0 & 1 \end{bmatrix} \tag{1}$$

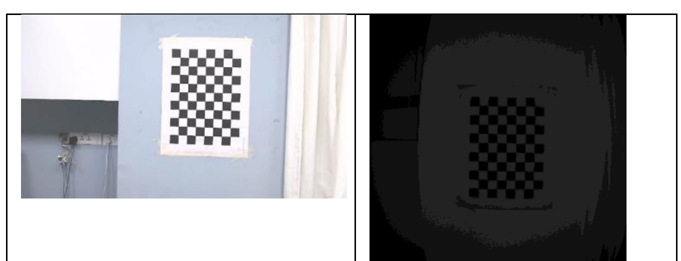
The ToF camera calibration was performed to estimate the intrinsic parameters of the ToF camera. The estimated calibration matrix is shown below.

$$K_{ToF} = \begin{bmatrix} fx_{tof} & 0 & Cx_{tof} \\ 0 & fy_{tof} & Cy_{tof} \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 388.88 & 0 & 262.63 \\ 0 & 388.76 & 254.86 \\ 0 & 0 & 1 \end{bmatrix} \tag{2}$$

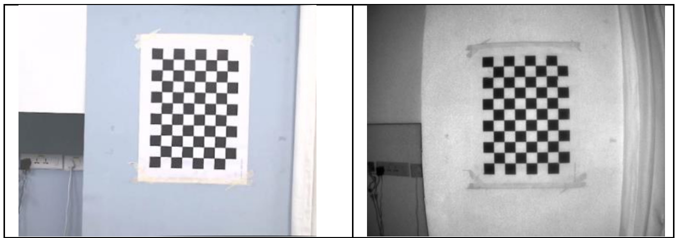
The scaling factor S to resize RGB image is calculated using the formula below:

$$\boldsymbol{S} = \frac{fx_{tof}}{fx_{rgb}} = \frac{388.88}{707.56} = 0.5496 \tag{3}$$

To align the resolution and pixel size of RGB and IR images, the RGB image is resized using a scale factor of 0.5496 resulting in a new image with a new resolution. The horizontal and vertical resolutions of RGB and IR images are compared. The image with the lower resolution determines the cropping dimensions for the other image. The higher-resolution image is cropped to match the lower-resolution image, ensuring both have the same resolutions for accurate calibration. The IR image captured by the ToF camera is then cropped to this resolution around its center. This ensures that both images share the same resolution and pixel size, facilitating accurate calibration for further processing. The original RGB image, original IR image, resized RGB image, and crop depth image are shown in Fig. 5.



Original RGB image (1280x720) | Original IR image (512x512)



Resized and cropped RGB image (512x396)	Cropped IR image (512x396)

Fig. 5. The Original RGB image, Original IR image, resized RGB image and crop depth image

TABLE I. COMPARISON OF CAMERA MATRIX AND REPROJECTION ERROR OF TOF CAMERA BEFORE AND AFTER IMAGE ENHANCEMENT

State	Camera Matrix	Reprojection Error
Before Enhancement	$\begin{bmatrix} 411.25 & 0 & 258.86 \\ 0 & 411.20 & 195.73 \\ 0 & 0 & 1 \end{bmatrix}$	0.8849
After Enhancement	$\begin{bmatrix} 392.09 & 0 & 262.09 \\ 0 & 392.07 & 194.34 \\ 0 & 0 & 1 \end{bmatrix}$	**0.3260**

The ToF camera calibration is performed first by employing IR image enhancement and then performing the steps for calibration as described in the methodology section, Intrinsic parameters of the ToF camera as well as reprojection error are estimated before and after IR image enhancement, which are shown in Table I. After acquiring the parameters from the single camera calibration of the ToF camera, the stereo calibration of ToF-RGB camera is performed to obtain the extrinsic parameters and reprojection errors, which are shown in Table II. During both single camera calibration and stereo camera calibration, reprojection error plays a significant role in authenticating the accuracy of the calibration process.

TABLE II. COMPARISON OF EXTRINSIC PARAMETERS AND REPROJECTION ERROR OF TOF-RGB STEREO CAMERA BEFORE AND AFTER ENHANCEMENT

State	Extrinsic Parameters	Error
Before	$\begin{bmatrix} 0.99899 & 0.04466 & -0.00386 & -78.267 \\ -0.04457 & 0.99877 & 0.02131 & 2.3388 \\ 0.00481 & -0.02112 & 0.99976 & 54.152 \end{bmatrix}$	0.75426
After	$\begin{bmatrix} 0.99897 & 0.04475 & -0.00659 & -83.9318 \\ -0.04454 & 0.99860 & 0.02835 & -1.0886 \\ 0.00785 & -0.02802 & 0.99957 & 3.9149 \end{bmatrix}$	**0.39394**

Table I, Table II, and Fig. 6 shows statistical analysis of the reprojection error before and after image enhancement during single and stereo camera calibration. Fig. 6 clearly shows that reprojection error is higher before enhancement compared to the reprojection error observed after enhancement. The enhancement of IR images improves corner detection, which helps to estimate more accurate calibration parameters of the camera and enhances the reliability of tasks such as undistortion, registration, and stereo matching.

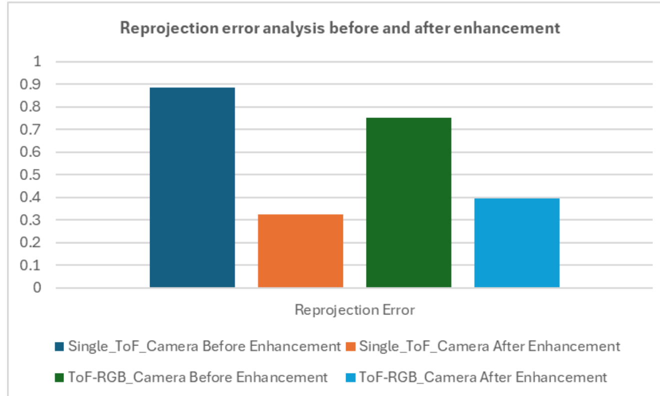


Fig. 6. Analysis of reprojection error of single and ToF-RGB camera before and after image enhancement.

B. *Comparative Study and Performance Analysis of Individual RGB, Individual ToF and ToF-RGB Camera Calibration Using MATLAB and OpenCV*

In this section, the performance evaluation was carried out to compare the accuracy and efficiency of individual camera and ToF-RGB camera calibration methods implemented in MATLAB and OpenCV. The collected datasets of a single RGB camera, a single ToF camera and ToF-RGB stereo camera have been tested with MATLAB and OpenCV. The intrinsic parameters such as focal length, camera center, distortion coefficients as well as reprojection error of RGB camera, ToF camera and TOF-RGB stereo camera have been computed using MATLAB and OpenCV library. Table III and Table IV show the camera matrix and distortion coefficients of both RGB and ToF cameras using MATLAB and OpenCV library. Table V and Fig. 7 represent the reprojection error of RGB camera, ToF camera and ToF-RGB stereo camera using MATLAB and OpenCV.

TABLE III. CAMERA MATRIX OF BOTH RGB AND ADI TOF CAMERAS USING MATLAB AND OPENCV LIBRARIES

Library	Sensors	Parameters			
		F_x	F_y	C_x	C_y
MATLAB	RGB Camera	394.02	394.33	253.68	187.83
	ToF Camera	389.59	388.99	262.78	200.82
OpenCV	RGB Camera	386.48	386.68	254.45	188.24
	ToF Camera	385.04	384.83	260.56	197.24

TABLE IV. DISTORTION COEFFICIENTS OF BOTH RGB AND TOF CAMERAS USING MATLAB AND OPENCV DIFFERENT LIBRARIES.

Library	Sensors	Parameters				
		K_1	K_2	P_1	P_2	K_3

MATLAB	RGB Camera	0.028	-0.194	-0.012	0.0014	0.0000
	ToF Camera	-0.159	-0.158	0.0011	-0.0000	0.0000
OpenCV	RGB Camera	0.0325	-0.156	-0.005	0.0005	0.2855
	ToF Camera	-0.128	-0.510	-0.001	0.00087	1.1674

TABLE V. REPROJECTION ERROR OF RGB CAMERA, TOF CAMERA AND TOF-RGB CAMERA USING MATLAB AND OPENCV LIBRARY

Sensors	MATLAB	OpenCV
RGB Camera	0.1491	0.221358
ToF Camera	0.2470	0.33850
Stereo calibration (RGB-ToF)	0.3781	0.468982

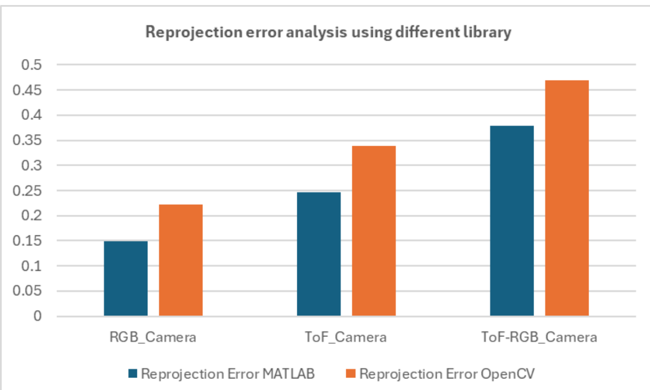


Fig. 7. Comparative analysis of reprojection error using MATLAB and OpenCV library

This study analyzed the calibration accuracy and performance of individual and ToF-RGB camera systems using MATLAB and OpenCV. By analyzing reprojection errors and comparing the calibration outcomes from both the MATLAB and OpenCV libraries, we observed variations in the camera matrix, distortion coefficients, and reprojection error. Fig. 7 and Table V shows that the reprojection error was found to be less than 0.5 with both MATLAB and OpenCV library. The reprojection errors of the RGB camera, ToF camera and ToF-RGB camera estimated using MATLAB were found to be less compared to reprojection errors estimated using OpenCV. During experimentation, we observed that MATLAB was able to detect corners in the checkboard at longer distance compared to OpenCV.

V. CONCLUSION

This paper presents a simple, reliable, and accurate approach to calibrate a ToF-RGB-like camera. The work proposes an IR enhancement approach to improve the quality of IR images obtained by a ToF camera and presents an approach to align the image resolution and pixel size of both the RGB and IR images. Additionally, steps for accurate calibration have been taken into consideration to find the intrinsic and extrinsic parameters of ToF-RGB camera. The effectiveness of the proposed approach has been evaluated based on reprojection error, before and after optimization. Furthermore, the calibrated images have been tested in other calibration libraries, including MATLAB and OpenCV. The experimental results demonstrate that the proposed method efficiently and precisely calibrates heterogeneous ToF-RGB camera and provides a robust foundation for different robotics applications.

REFERENCES

[1] S. Fuchs and G. Hirzinger, "Extrinsic and depth calibration of ToF-cameras," In 2008 IEEE Conference on Computer Vision and Pattern Recognition, pp. 1-6. IEEE, 2008.

[2] K. Konolige, "Projected texture stereo," In 2010 IEEE International Conference on Robotics and Automation, pp. 148-155. IEEE, 2010.

[3] F. Endres, H. Jürgen, S. Jürgen, C. Daniel, and W. Burgard, "3-D mapping with an RGB-D camera," IEEE transactions on robotics, vol. 30, no. 1 pp. 177-187, 2013.

[4] C.J. Taylor and A. Cowley, "Parsing indoor scenes using rgb-d imagery," 2013.

[5] M. Munaro and E. Menegatti, "Fast RGB-D people tracking for service robots," Autonomous Robots, vol. 37, pp. 227-242, 2014.

[6] J. Tang, S. Miller, A. Singh, and P. Abbeel, 2012, May. "A textured object recognition pipeline for color and depth image data," In 2012 IEEE International Conference on Robotics and Automation, pp.3467-3474, IEEE, 2012.

[7] A. Canessa, M. Chessa, A. Gibaldi, S.P. Sabatini and F. Solari, "Calibrated depth and color cameras for accurate 3D interaction in a stereoscopic augmented reality environment," Journal of Visual Communication and Image Representation, vol. 25, no. 1, pp.227-237, 2014.

[8] Y. Cui, S. Schuon, D. Chan, S. Thrun and C. Theobalt, "3D shape scanning with a time-of-flight camera," In 2010 IEEE computer society conference on computer vision and pattern recognition, pp. 1173-1180. IEEE, 2010.

[9] A. Ramey, V. Gonzalez-Pacheco and M.A. Salichs, 2011, "Integration of a low-cost RGB-D sensor in a social robot for gesture recognition," In Proceedings of the 6th international conference on Human-robot interaction, pp. 229-230, 2011.

[10] B. Huang and S. Zou, "A new camera calibration technique for serious distortion," Processes, vol. 10, no. 3, pp.488, 2022.

[11] Y. Zhou, D. Chen, J. Wu, M. Huang and Y. Weng, "Calibration of RGB-D camera using depth correction model," In Journal of Physics: Conference Series, vol. 2203, no. 1, p. 012032, 2022.

[12] H. Liu, H. Li, X. Liu, J. Luo, S. Xie and Y. Sun, 2019. "A novel method for extrinsic calibration of multiple RGB-D cameras using descriptor-based patterns," Sensors, vol. 19, no. 2, p.349, 2019.

[13] P.C. Su, J. Shen, W. Xu, S.C.S. Cheung, and Y. Luo, 2018. "A fast and robust extrinsic calibration for RGB-D camera networks," Sensors, vol. 18, no. 1, p.235, 2018.

[14] A. Perez-Yus, E. Fernandez-Moral, G. Lopez-Nicolas, J.J. Guerrero and P. Rives, "Extrinsic calibration of multiple RGB-D cameras from line observations," IEEE Robotics and Automation Letters, vol. 3, no. 1, pp.273-280, 2017.

[15] G. Chen, G. Cui, Z. Jin, F. Wu and X. Chen, "Accurate intrinsic and extrinsic calibration of RGB-D cameras with GP-based depth correction," IEEE Sensors Journal, vol. 19, no. 7, pp.2685-2694, 2018.

[16] F. Basso, E. Menegatti and A. Pretto, "Robust intrinsic and extrinsic calibration of RGB-D cameras," IEEE Transactions on Robotics, vol. 34, no. 5, pp.1315-1332, 2018.

[17] W. Darwish, S. Tang, W. Li and W. Chen, "A new calibration method for commercial RGB-D sensors," Sensors, vol. 17, no. 6, p.1204, 2017.

[18] O, Semeniuta, "Analysis of camera calibration with respect to measurement accuracy" Procedia Cirp, vol. 41, pp.765-770, 2016.

[19] A.N. Staranowicz, G.R. Brown, F. Morbidi and G.L Mariottini, "Practical and accurate calibration of RGB-D cameras using spheres," Computer Vision and Image Understanding, vol. 137, pp.102-114, 2015.

[20] Analog Devices, "ToF ROS2," GitHub, [Online]. Available: https://github.com/analogdevicesinc/tof-ros2/blob/main/src/aditof_utils.cpp

[21] Intel, "D400 Series Custom Calibration White Paper," Intel RealSense, [Online]. Available: https://dev.intelrealsense.com/docs/d400-series-custom-calibration-white-paper.