A Benchmark Dataset for RGB-D Sphere Based Calibration

David J.T. Boas *,[1,2], **Sergii Poltaretskyi** [2], **Jean-Yves Ramel** [1], **Jean Chaoui** [2], **Julien Berhouet** [1] , **Mohamed Slimane** [1]

[1]Laboratoire d'Informatique Fondamentale et Appliquée de Tours (LIFAT), Université François Rabelais, 37200 Tours, France

[2]IMASCAP, 29280, Plouzané, France

* david.boas@etu.univ-tours.fr

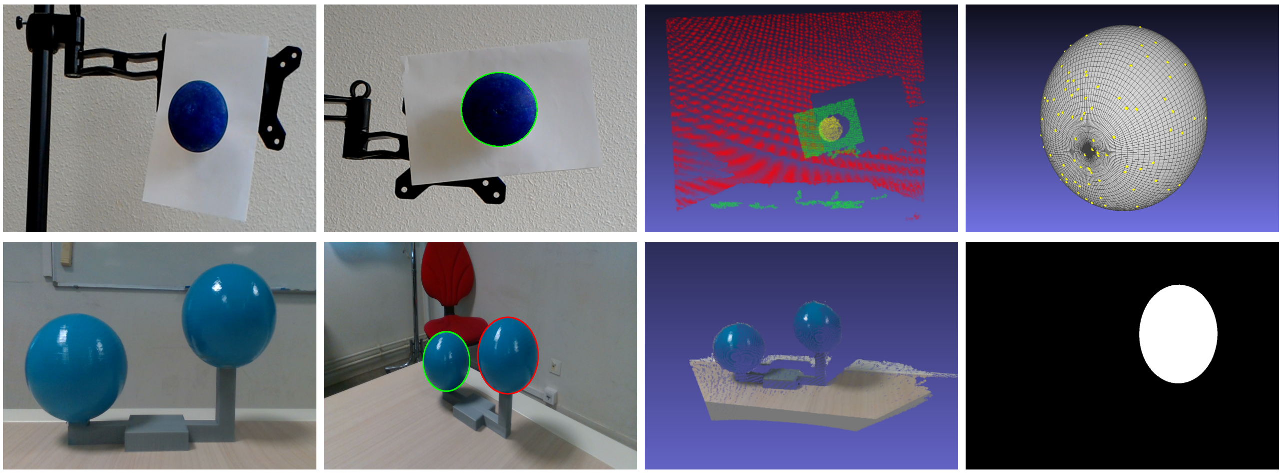


Figure 1: RGB and Depth Images of Real and Synthetic Dataset (fourth column) in our benchmark dataset for RGB-D sphere based calibration. The proposed dataset contains 3 real sequences and 1 synthetic sequence with varying noise. The dataset contains all camera parameters, as well as pre-detected ellipses and spheres.

Keywords: RGB-D Calibration, Sphere based, Dataset, Benchmark

Abstract

Accurate calibration of an RGB-D camera couple is an important task in computer vision, especially applications requiring the knowledge of 3D information. Although a variety of algorithms have been proposed, it remains difficult to evaluate existing methods in the literature as oftentimes different sequences are used. In this paper, we propose a full dataset benchmark, with real and synthetically generated sequences, manually determined ground truth, and evaluation metrics for comparison. Evaluation of three methods using this framework are also provided. The proposed benchmark dataset is available online at http://www.rfai.li.univ-tours.fr/tools-and-demos/.

1 Introduction

A digital representation of our environment is a fundamental requirement in many computer vision applications. An accurate camera calibration is essential in many of these algorithms, such as 3D reconstruction, object tracking or camera localization. A key element of a more accurate representation of the observed scene is to determine the three-dimensional geometry of the objects of interest. However, the process of image formation loses the depth information. Recently, the availability of low cost depth cameras has made it possible to easily capture this 3D data. The combination of a depth sensor with an RGB camera provides enhanced scene information, and is referred to as an RGB-D device. To obtain reliable data, it is necessary to model the specific properties of the sensor pair with a calibration procedure (Figure 2). However, the quality of the manufacturer's calibration, as well as the sensitivity of the calibration to external conditions, often makes it necessary to perform this operation by an expert, especially for high-precision applications. (like medical ones). Also, these parameters may not be given by the constructors.

To the authors knowledge, there is no general sphere based RGB-D calibration dataset publicly available. New approaches defines their own datasets, but the use of different cameras and calibration scenarios makes the comparison between several algorithms difficult. In this work, we propose a publicly available benchmark dataset to evaluate sphere based calibration, with both real and synthetically generated data, as well as results from three methods [1, 2, 3].

This paper is structured as follows. Section 2 surveys recent works related to RGB-D sphere based calibration. Section

3 detail how the dataset was built, as well as several design choices. In Section 4, multiples metrics are proposed to assess the quality of an RGB-D calibration. Finally, Section 5 conclude this paper.

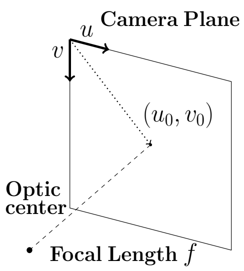
2 Related work

Most calibration approaches use objects whose geometric properties are known, and are said to be supervised. These approaches are based on the principle of identifying matching features observed from several points of view to determine the calibration parameters. The sphere appears as a points cloud for the depth camera, and as an ellipse for the RGB camera. The estimated centers of these quadratic forms provide reference points for calibration (Figure 3). Sphere based approaches are easy to set up, as such an object is visible regardless of the point of view of the cameras. In addition, their center can be reliably detected by both cameras. Its disadvantage lies in the fact that only a few calibration points are captured per frame, making them more sensitive to measurements errors.

RGB Sphere based calibration Spheres are also a commonly used reference object in RGB calibration approaches. These methods either uses the ellipses outline to deduce tangent lines as calibration features [4], or the knowledge of the sphere radius to estimate the sphere's center in 3D as a reference point [5, 6]. Their calibration approaches reaches accuracies similar to reference RGB checkerboard based calibration [7].

RGB-D Calibration Similarly to RGB camera calibration, some RGB-D approaches uses a checkerboard as a calibration reference object. The core idea is to identify checkerboard corner points in the infrared image[8, 9], and to use the knowledge that the checkerboard lies on a plane [10, 11]. However, these methods are not ideal, as the user must manually move the checkerboard, which has to be printed in large format to be visible by both cameras (especially with a wide baseline).

RGB-D Sphere based calibration Spheres have proven to be a reliable calibration object for RGB-D camera calibration. The method of Staranowicz et al.[2] proposes the use of a single sphere with unknown geometric properties, and to correct the projected spheres centers. However, they do not model the Depth mapping function. Several sphere based approaches are based on the principle of precisely determining the center of the sphere, and to register the spheres centers identified by both



$$K := \begin{pmatrix} f_u & 0 & u_0 \\ 0 & f_v & v_0 \\ 0 & 0 & 1 \end{pmatrix}$$

Figure 2: Visualization of the intrinsic parameters of the pinhole camera model.

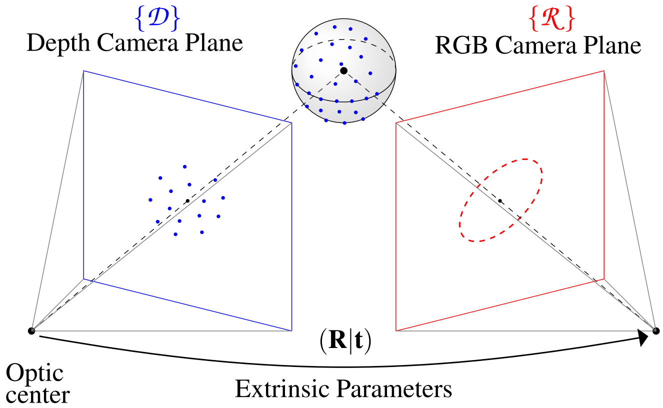


Figure 3: Overview of a sphere based RGB-D calibration scene. The sphere is seen as an ellipse for the RGB camera {R}. The sphere is seen as a point cloud for the depth Camera {D}. The two optic centers are linked by a rigid transform $(\mathbf{R}|\mathbf{t})$.

cameras. Using the knowledge of the sphere radius, Boas et al. [3] proposed to detect spheres centers from RGB images and to perform the calibration using the 3D spheres centers coordinates, which limits the effect of the depth camera error increasing with the distance. Another advantage of sphere based calibration is to allow calibration with wide baselines, making it effective for camera network extrinsic parameters estimation [12, 13, 14]. However, these approaches rely mostly on the depth data, and only estimate the extrinsic parameters.

Datasets Previous methods define their own dataset, which makes it difficult to repeat some experiments and to make comparisons on an identical basis. Indeed, a calibration algorithm is very much linked to the detection phase, as well as to the hardware used. We thus propose a common basis for the evaluation of sphere based calibration.

3 Proposed benchmark dataset

A first step for sphere based RGB-D calibration is to detect the spheres precisely. However, a common problem for camera calibration evaluation is the impossibility of absolutely determining both the intrinsic parameters of the cameras, and the extrinsic parameters between the two cameras. Thus, without reference values, it is difficult to assess camera calibration quality. The usual way to get around this problem is to define, in addition to real data, synthetically generated data whose parameters are known. Thereby, our dataset is divided into real and synthetic data, with ground truth determined manually for real data. An overview of all proposed dataset is presented in Table 1.

3.1 Problem formulation

To align the RGB and Depth streams, it is necessary to model the intrinsic parameters of both cameras. In the case of a standard camera, these parameters are represented by a pinhole camera model, which models the image formation process by

Dataset	Color Camera Resolution	Depth Camera Resolution	# RGB-D Pairs	Support	Capture Setup	Known Parameters
Random	Hololens n°1 1280x720	Structure Sensor 640x480	72	Single Sphere Support	Stationary	Intrinsic RGB Intrinsic Depth
Equidistant	Hololens n°2 1280x720	Structure Sensor 640x480	27	Single Sphere Support	Stationary	Intrinsic RGB Intrinsic Depth
Double	Realsense SR300 1920x1080	Realsense SR300 640x480	40	Double Sphere Support	Hand-held	Intrinsic RGB Intrinsic Depth Extrinsic RGB-D
Synthetical	1280x960	-	20	Single sphere	-	All (Absolute values)

Table 1: Proposed datasets overview for sphere based RGB-D calibration.

representing the focal distance and the center of the image (also known as the principal point) into a 3x3 matrix.

The intrinsic parameters $\mathbf{K}$ of the two cameras are referred to as ${}^{\mathcal{R}}\mathbf{K}$ and ${}^{\mathcal{D}}\mathbf{K}$, where $\{\mathcal{R}\}$ refers to the RGB camera and $\{\mathcal{D}\}$ to the depth camera; with (f_u, f_v) the focal length and (u_0, v_0) the principal point. The alignment of the two streams is then achieved by a rigid transform $(\mathbf{R}|\mathbf{t})$, called the extrinsic parameters. This transform is composed of a 3x3 rotation matrix $\mathbf{R}$ and a 3x1 translation vector $\mathbf{t}$. The rotation can also be modeled by a rotation vector $\mathbf{r} := (r_x, r_y, r_z)$, which is expressed as Euler angles in degrees (ZYX convention).

3.2 Real sequences

3.2.1 Acquisition

While acquiring a sequence, the two cameras simultaneously observe the same scene. These two cameras are placed as close as possible, and are considered motionless relative to each other. Both cameras observe a sphere whose radius is known. These cameras capture RGB-D image pair (i.e. an RGB image and a Depth point cloud) simultaneously, which are available for each sequence. Each image pair is temporally distinct (for improved quality), and was captured in good and fixed luminosity conditions.

To construct the described datasets, we used 3D printed spheres (printing accuracy of 0.5 *mm*). To simplify the ellipse detection phase, these spheres are painted in blue. Two types of support have been created as part of this benchmark (Figure 4).

Single Sphere support The first support consist of a single sphere attached to a mobile mast. The mast is fixed to a table. A sheet of paper is added behind the sphere to give a high contrast between the sphere and the background. While making captures, a human operator manually move the sphere to its next position.

The data are captured with the following cameras. For RGB images, the front facing color camera of the Microsoft mixed reality headset, the Hololens 1.0, is used. A structured light depth camera, the Structure Sensor, is attached above the headset. The headset is placed on a fixed 3D printed support, and is oriented towards the sphere. The manufacturers intrinsic camera parameters are known. However, no manufacturer extrinsic parameters can be provided. Two sequences following this setup are provided.

Random For this sequence, the sphere is placed randomly in the field of view of both cameras, for a total of 72 RGB-D image pairs. The relative distance between the sphere and the cameras for a capture is comprised between 0.35 and 0.75 meters.

Equidistant This sequence, on the contrary, controls precisely the sphere position. As such, the positions of the spheres are determined in such a way as to uniformly cover the entire field of view of the cameras. For a given depth value, nine positions are determined. Eight of these positions follow the edge of the RGB camera's field of view. The last position is the center of the camera. This operation is performed for three distinct depth values, i.e. 0.5, 0.75 and 1 meters, providing 27 RGB-D image pairs in total.

Double Sphere support The second support consists of two spheres placed on a stable support, all entirely printed in 3D. For the related sequence, named **Double**, the support is fixed, and the camera is hand-held. The camera couple used here is the Intel RealSense SR300. This sequence allows for a more realistic setup, and to capture a wider number of sphere by RGB-D image pairs.

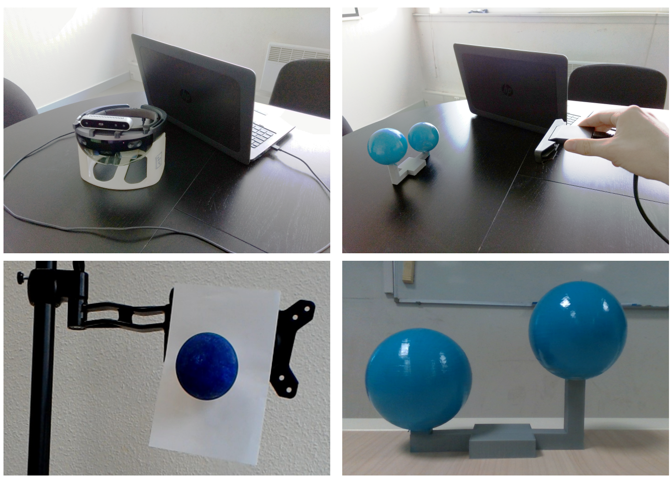


Figure 4: Visualization of both capture setups, as well as single and double sphere support.

In total, 20 RGB-D image pairs were captured, giving a number of 40 spheres. Finally, both the intrinsic and extrinsic parameters are provided by the manufacturer.

3.2.2 Ground truth

There is no mean to have an accurate estimate of the intrinsic and extrinsic parameters with real data other than performing a calibration. Nonetheless, to have a comparison value, input RGB-D image pairs have been manually segmented with high accuracy. The Direct Linear Transform (DLT) algorithm, adapted to sphere based calibration [1, 2], have been applied in order to have an estimate of the intrinsic and extrinsic parameters. This algorithm requires both ellipses detection for the RGB camera, and spheres detection for the depth camera.

RGB Ellipse detection For the related sequences, the ellipses have been manually segmented using the Gimp Software. The contour is then extracted, and an optimization based least-square ellipse fitting is performed.

For each RGB frame, we thus obtain an ellipse estimate, as well as a binary mask. This mask is used to estimate ellipse detection algorithm's accuracy against ground truth with a pixel-by-pixel comparison.

Depth Sphere detection Regarding the depth data, the points cloud have been manually segmented with the MeshLab Software. Once only points belonging to the sphere remains, an optimization based least-square sphere fitting is performed, using the sphere radius knowledge. The cropped points are projected back onto the depth camera plane using the manufacturer depth intrinsic parameters, giving a segmented depth map, in a matrix form. Thus, for each depth frame, we obtain a cropped point cloud estimate, a sphere estimate and a depth map estimate.
Each type of output data, as well as their format, are explained in detail in subsection 3.4.

3.3 Synthetical sequences

3.3.1 Generation

While the accuracy of the detection algorithms used as the first step of every calibration algorithm may be improved, an error will always remain because of various factors (rasterization, sensor noise, analogical to digital conversion, etc ...). It is well-known that calibration algorithms are sensitive, and highly dependent on the detection accuracy of the reference points. It is thus common to evaluate any new calibration algorithm with synthetically generated data, allowing perfect prior knowledge of both extrinsic and intrinsic parameters.

The simulated scene contains an RGB-D camera couple observing multiples spheres with known radius. The sphere's positions, the extrinsic parameters and the intrinsic parameters of both cameras are known and fixed for all synthetically generated dataset, and will constitute our ground truth. The depth camera is placed above the color camera at a reasonable distance, and with no rotation. The intrinsic parameters are chosen with values according to currently available hardware.

As proposed by [2], twenty spheres are randomly generated at the intersection of the field of view of the two cameras. For each sphere center, we randomly generate a hundred 3D points on the periphery of half of the sphere by using the parametric equation of a sphere. We generate the input data from these twenty sets of a hundred points. Knowing the true properties of the spheres, we can project the spheres on the RGB camera plane to obtain the associated ellipses. Also, we generate the depth map of the depth camera by projecting the points on the depth camera plane, in a matrix form.

3.3.2 Noise application

Once the reference ground truth scene is generated, some noise has to be introduced to the data to simulate a realistic scene. Three kind of errors are introduced, with values selected by reviewing recent RGB camera calibration literature [7, 15].

RGB ellipse fitting error The error f_1 models the image formation error and the ellipse fitting error. It applied a zero mean Gaussian noise $\mathcal{N}(0, \sigma_{\mathcal{R}} = 0.6)$ on all points belonging to the detected ellipses. The ellipse fitting process is then applied back on these modified points. This method allows to change the ellipses centers, their axes length, and their orientation consistently.

RGB intrinsic noise The second error f_2 models the inaccuracies in the RGB camera intrinsic parameters estimates. It is represented by Equation 1, which applies a linear noise on the intrinsic parameter by scaling their value by a scalar $\eta = 1 + 2‰$.

$$f_2(K, \eta) = \begin{pmatrix} \eta * f_u & 0 & \eta * u_0 \\ 0 & \eta * f_v & \eta * v_0 \\ 0 & 0 & 1 \end{pmatrix} \tag{1}$$

Depth displacement The third error, σ_D, represents the error in sphere fitting, as well as depth camera inaccuracies increasing with distance. It is modeled by an identical translation on all sphere points, which has the effect of moving the detected sphere center. This noise allows to characterize any kind of error. Indeed, a Gaussian noise only slightly changes the points cloud centroid, making this approach inefficient.

We wish to evaluate the influence of the estimated sphere centers on the calibration result. Thus, we only vary the Depth displacement σ_D. This noise is randomly generated for each sphere's points cloud. The random aspect of this noise requires to repeat the experiment several times (here 100 times) to obtain average results.

3.4 Datasets organization

For every dataset, raw data captured by the sensors, manually processed data (as ground truth) and automatically processed data are available. Three kind of data format are supplied. Images files are provided as **.PNG**, points cloud files as **.PLY**, and parameters files are **.YML**. The provided files are separated in several folders, with associated files having the same filename (Figure 5). Finally, we also propose the manufacturer calibration parameters, as well as the sphere radius.

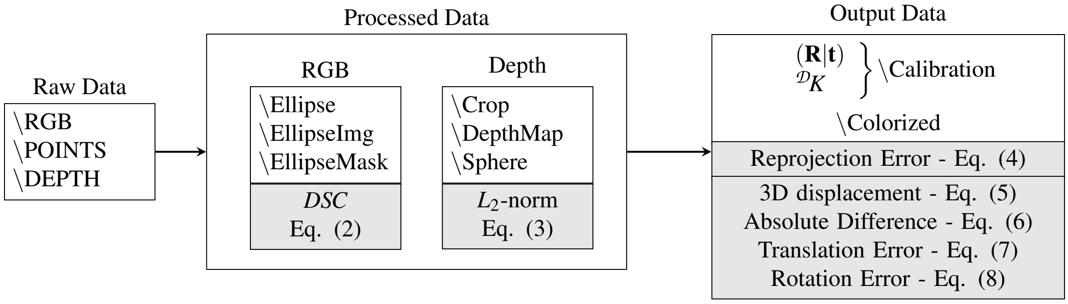


Figure 5: Overview of dataset organization and their respective metrics (in light gray) for evaluation of the RGB-D Calibration process.

Raw Data Data provided by the sensors are separated in three folders. We provide color images, point clouds, as well as depth maps obtained by the cameras. No sensor depth maps are provided for the Double dataset, and the synthetically generated one.

Processed Data Processed Data include segmented files, as well as fitting results of the identified quadratic forms. This allows to begin, and evaluate, a calibration algorithm at any point of the calibration pipeline.

RGB related processed data are linked to ellipses. Those are provided in parametric form. The ellipses masks, as well as their visualization on the RGB images, are provided.

Depth related processed data include segmented points cloud, associated segmented depth maps, and fitted spheres in parametric form. We also supply the projection of the sphere centers onto the depth camera plane, and their fitting error (Root Mean Square).

Output Data In addition, we give calibration results from the DLT algorithm [1]. Using the DLT results as an initial calibration guess, we also propose results for the Staranowicz et al. method [2], and Boas et al. approach [3]. All following evaluation metrics, as well as colorized points cloud are also available. Finally, an analysis of the calibration results has been provided in [3].

4 Evaluation Methodology

In our work, we propose several metrics to evaluate the calibration accuracy. We will refer to synthetically generated calibration parameters, as well as manually segmented data as our ground truth GT.

4.1 Ellipse detection evaluation

To evaluate ellipse detection, a comparison pixel-by-pixel is proposed. The pixels on the outline and inside the estimated ellipse, X, are compared against ground truth binary mask pixels X_{GT}. A single evaluation value is given by the Sørensen-Dice coefficient (Equation 2).

$$DSC = \frac{2|X \cap X_{GT}|}{|X| + |X_{GT}|} \tag{2}$$

4.2 Sphere detection evaluation

To evaluate Depth data segmentation, we jointly evaluate the sphere detection and fitting. We use the L_2-norm (also known as the Euclidean distance) between the manually estimated sphere center ${}^{\mathcal{D}}\mathbf{O}_{GT}$, and the obtained sphere center ${}^{\mathcal{D}}\mathbf{O}$ (Equation 3).

$$L_2 = ||{}^{\mathcal{D}}\mathbf{O}_{GT} - {}^{\mathcal{D}}\mathbf{O}||_2 \tag{3}$$

4.3 Calibration evaluation

4.3.1 Real datasets metrics

Two evaluation metrics for real data are proposed : the gold standard 2D reprojection error [1], and the visualization.

Reprojection error It is the mean Euclidean distance between the estimated RGB ellipses centers,${}^{\mathcal{R}}\mathbf{O}_{e_i}$, and the projection of the depth sphere center ${}^{\mathcal{D}}\mathbf{O}_{s_i}$ onto the color camera plane (Equation 4). This computation requires the rigid transform $(\mathbf{R}|\mathbf{t})$ between the depth and RGB camera to express the sphere center in the RGB camera coordinate system $({}^{\mathcal{R}}x_i, {}^{\mathcal{R}}y_i, {}^{\mathcal{R}}z_i)$. The point is then projected using the RGB camera intrinsic parameters ${}^{\mathcal{R}}\mathbf{K}$.

$$e_r = \frac{1}{n}\sum_{i=1}^{n} ||{}^{\mathcal{R}}\mathbf{O}_{e_i} - \frac{1}{{}^{\mathcal{R}}z_i}\, {}^{\mathcal{R}}\mathbf{K}(\mathbf{R}\, {}^{\mathcal{D}}\mathbf{O}_{s_i} + \mathbf{t})||_2 \tag{4}$$

Visualization Using the same approach to project depth data onto the RGB camera plane, it is possible to obtain colorized points cloud. This offers a good visualization of the calibration parameters integrity, especially on objects borders.

4.3.2 Synthetic Dataset

Although the reprojection error allows to evaluate the calibration with a single value, it does not allow to determine the cause of a calibration error. The knowledge of real intrinsic and extrinsic parameters with synthetically generated data allows to define several more evaluation metrics to identify the error origin.

3D displacement This metric can be interpreted as the 3D equivalent of the reprojection error. It allows to express all calibration parameters in a single value, while avoiding the problems of the reprojection error. It is the mean Euclidean distance between the ground truth RGB spheres centers ${}^{\mathcal{R}}\mathbf{O}_{s_i}$ and their corresponding depth spheres centers ${}^{\mathcal{D}}\mathbf{O}_{s_i}$ (Equation 5). This metric cannot be used with real data, as it can be difficult to accurately detect the center of a sphere from a single RGB image.

$$E_r = \frac{1}{n}\sum_{i=1}^{n} ||{}^{\mathcal{R}}\mathbf{O}_{s_i} - {}^{\mathcal{R}}\mathbf{K}(\mathbf{R}\ {}^{\mathcal{D}}\mathbf{O}_{s_i} + \mathbf{t})||_2 \tag{5}$$

Absolute difference We also propose to use the absolute difference between a real value X_{GT} and the estimated value X (Equation 6).

$$X_{diff} = |X - X_{GT}| \tag{6}$$

This allows to determine the error for any parameter included in the calibration, i.e. (f_u, f_v, u_0, v_0) for the intrinsic parameters, and $(r_x, r_y, r_z, t_x, t_y, t_z)$ for the extrinsic ones.

Rigid body errors An easier way to assess the extrinsic parameters accuracy in fewer values is to evaluate both the rotation and the translation errors in their entirety. Thus, we propose to use the L_2-norm to evaluate the translation error $\mathbf{t}_{err}$ (Equation 7).

$$\mathbf{t}_{err} = ||\mathbf{t}_{GT} - \mathbf{t}||_2 \tag{7}$$

As for the rotation, a direct way to evaluate rotation matrices is to perform the dot product between the ground truth $\mathbf{R}_{GT}$ and the estimate $\mathbf{R}$ to have the relative rotation. The trace of the resulting rotation matrix makes it possible to directly determine the value of its rotation angle θ, which is equal to $1 + 2cos(\theta)$. Thus, the relative rotation value, which corresponds to the error $\mathbf{R}_{err}$, is expressed by Equation 8 in radians.

$$\mathbf{R}_{err} = acos(\frac{trace(\mathbf{R}_{GT}^T.\mathbf{R}) - 1}{2}) \tag{8}$$

5 Conclusion

In this work, we propose a benchmark dataset for sphere based RGB-D calibration, and propose metrics at every step of a calibration process to assess their performance. The proposed benchmark dataset contains 3 real datasets, and one synthetically generated dataset, with respective manually obtained ground truth for real data, and absolute ground truth for synthetic data. The evaluation results of three approaches are also provided.

References

[1] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN: 0521540518, second ed., 2004.

[2] A. Staranowicz, G. R. Brown, F. Morbidi, and G. L. Mariottini, "Easy-to-Use and Accurate Calibration of RGB-D Cameras from Spheres," in *Image and Video Technology*, vol. 8333, pp. 265–278, Springer Berlin Heidelberg, 2014.

[3] D. J. T. Boas, S. Poltaretskyi, J.-Y. Ramel, J. Chaoui, J. Berhouet, and M. Slimane, "Relative pose improvement of sphere based rgb-d calibration," *Proceedings of the 14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, vol. 4, pp. 91–98, 2019.

[4] K.-Y. K. Wong, G. Zhang, and Z. Chen, "A Stratified Approach for Camera Calibration Using Spheres," *IEEE Transactions on Image Processing*, vol. 20, no. 2, pp. 305–316, 2011.

[5] D. Schnieders and K.-Y. K. Wong, "Camera and light calibration from reflections on a sphere," *Computer Vision and Image Understanding*, vol. 117, pp. 1536–1547, 2013.

[6] J. Sun, H. He, and D. Zeng, "Global Calibration of Multiple Cameras Based on Sphere Targets," *Sensors*, vol. 16, no. 1, p. 77, 2016.

[7] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 11, pp. 1330–1334, 2000.

[8] D. Herrera C., J. Kannala, and J. Heikkila, "Joint Depth and Color Camera Calibration with Distortion Correction," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 10, pp. 2058–2064, 2012.

[9] I. V. Mikhelson, P. G. Lee, A. V. Sahakian, Y. Wu, and A. K. Katsaggelos, "Automatic, fast, online calibration between depth and color cameras," *Journal of Visual Communication and Image Representation*, vol. 25, no. 1, pp. 218–226, 2014.

[10] 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.

[11] F. Basso, E. Menegatti, and A. Pretto, "Robust Intrinsic and Extrinsic Calibration of RGB-D Cameras," *IEEE Transactions on Robotics*, pp. 01–18, 2018.

[12] J. Shen, W. Xu, Y. Luo, P.-C. Su, and S.-c. S. Cheung, "Extrinsic calibration for wide-baseline RGB-D camera network," IEEE 16th International Workshop on Multimedia Signal Processing , 2014.

[13] Minghao Ruan and D. Huber, "Calibration of 3D Sensors Using a Spherical Target," *Proceedings of the International Conference on 3D Vision*, vol. 01, pp. 187–193, 2014.

[14] P.-C. Su, J. Shen, W. Xu, S.-C. Cheung, and Y. Luo, "A Fast and Robust Extrinsic Calibration for RGB-D Camera Networks," *Sensors*, vol. 18, no. 1, p. 235, 2018.

[15] Z. Liu, Q. Wu, S. Wu, and X. Pan, "Flexible and accurate camera calibration using grid spherical images," *Optics Express*, vol. 25, no. 13, pp. 15269–15285, 2017.