

Improvements of the Accuracy of Generic Extrinsic Camera Calibration with ICP

Carsten Schmerbeck¹ ; Linus Dau¹; Michael Heizmann¹ 

¹Karlsruhe Institute of Technology, Institute of Industrial Information Technology (IIIT), Karlsruhe

Abstract

For the extrinsic calibration of a camera system, a precise estimation of a set of calibration points is necessary [1]. However, there are influences that reduce the accuracy of point localization, such as lighting, exposure time, and thereby resulting in the over- and underexposure of the recorded images. This paper demonstrates the influence of varying exposure times on the localization of ArUco calibration markers [2] for generic ICP-based extrinsic calibration [1, 3, 4]. We evaluated the perceived location of the marker corners and thus their perceived size for various exposure times due to blur and overexposure. Furthermore, we propose a method of enabling subpixel accuracy by combining adjacent rays with a weighted mean. For all settings, we quantified the effect on the extrinsic camera calibration. We found, that the markers appear smaller with increasing exposure times, which can be explained by blur of the marker borders.

Keywords: Camera calibration, ICP

1 Introduction

1.1 Classic intrinsic camera calibration

The intrinsic camera calibration is the process to determine the internal camera parameters. A precise determination of these parameters is paramount for computer vision and optical metrology applications. The intrinsic parameters are specific for a camera and need to be calibrated for each camera individually. These parameters include the focal length (f_x, f_y) and the optical centers (c_x, c_y), which can be used to create the camera matrix $\mathbf{K}$. Furthermore, there are distortion coefficients that can improve the quality of the calibration [2].

$$\mathbf{K} = \begin{pmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{pmatrix} \quad (1)$$

1.2 Generic intrinsic camera calibration

For some applications, these few parameters are not enough to describe the camera sufficiently accurate. This finding is backed by the experiments of Schöps et al. [5]. This especially applies to non-central cameras such as plenoptic cameras [3].

Generic intrinsic calibration is a model-free black-box representation of a camera, consisting of a ray or raxel representation of each camera pixel individually, as introduced by Grossberg and Nayar [6].

1.3 Exposure in camera calibration

Illumination and exposure significantly influence the performance of camera calibration systems. This is backed by various publications, e.g.

- Cledat et al. [7] showed the change of perceived size of ArUco markers with varying exposure.

- Experiments of Mousavirad et al. [8] showed this relation for a distance estimation system.
- Agnus et al. [9] proposed an approach to estimate the effect of overexposure in rectangular ArUco patterns.
- Mondéjar-Guerra et al. [10] simulated the effect of overexposure on ArUco markers using a dilation operation.
- Berral-Soler et al. [11] proposed a deep learning-based framework for ArUco detection in challenging lighting conditions.

2 Hardware setup

2.1 Cameras

The cameras used in this experiment are made of Raspberry Pi 3 [12] and Raspberry Pi HQ Camera modules with C/CS-mount. These parts are built into an aluminum tube, which grants the cameras the ability to be stacked or easily mounted into the calibration setups. The system is explained in more detail in [1, 13].



(a) Intrinsic setup.

(b) Extrinsic setup.

Figure 1: Hardware setups for calibration.

2.2 Intrinsic calibration setup

For the intrinsic calibration, a single camera is placed in front of a monitor, which serves as an active reference target. The camera is mounted on a linear axis to be able to be moved automatically towards the monitor. This setup is put in an opaque enclosure whose walls are painted black to absorb any stray light from the monitor. The system is explained in greater detail in [1, 13].

2.3 Extrinsic calibration setup

For the extrinsic calibration, the intrinsically calibrated cameras are mounted on a metal frame in varying positions directed at a reference plane below them. On this plane, ArUco boards are placed as reference targets. The locations of the ArUco boards are known and are used to estimate the camera's position and orientation [14].

This is achieved by placing the reference targets on a flat reference plane, a precision-milled piece of cast aluminum¹ measuring $1000 \times 1000 \times 25$ mm. The material has a specified surface flatness of $130 \mu\text{m m}^{-1}$, as well as extremely low internal stresses [15]. The plate is set up on six equally spaced, adjustable feet that were adjusted until no more deviations could be detected using a 1000 mm precision steel straightedge² with straightness according to DIN874 GG1 ($\approx 20.67 \mu\text{m}$) [16, p. 181] and a $20 \mu\text{m}$ feeler gauge³. This procedure of measurement is in accordance to DIN876:2 [17, Sec. 4.3]. A placement of the plate on its three Bessel points [17, Sec. 3.1][18, Sec. 3.1.5, 5.4.2] did not produce a sufficiently flat surface.

The positioning of the targets is achieved through aligning them with a plus-shaped 3 mm sheet of aluminum, that was milled around its entire circumference to the required contour within a specified tolerance of $\pm 50 \mu\text{m}$ ⁴. The targets themselves are 4 mm aluminum composite plates, on which the ArUco markers were printed⁵. Beforehand, the markers were positioned in a milling machine by the corners of the prints using a centering microscope⁶, and their circumference was milled to match the desired contour within a specified tolerance of $\pm 50 \mu\text{m}$ ⁷.

3 Generic calibration procedure

3.1 Generic intrinsic calibration

The intrinsic parameters were determined in two steps identical to the description by Zörn et al. [1]: (i) The monitor displays a sequence of sinusoidal patterns, whereby the phase of the signal contains

¹G.AL C250 by Gleich Aluminium GmbH

²0465103 by Helios Preisser / Mahr GmbH

³W2868B by Jiechao Tools Co., Ltd.

⁴Milled with CAM 1020 Premium by vhf camfacture AG

⁵Unidirectional UV direct printing on 4 mm aluminum composite board by Hannes Industriegravuren GmbH

⁶MZM30 by Erdmann & Grün KG Wetzlar

⁷Milled with FP2 by Friedrich Deckel AG, measurements with LS388C by Dr. Johannes Heidenhain GmbH

coding on the pixel coordinate. For each pattern displayed, an image is taken from the camera at 10 equidistant known positions in front of the monitor. The positioning was automated using a motorized linear axis, moving the camera towards the monitor [1]. Through probabilistic phase unwrapping the registration between the camera pixels and monitor is determined. For this, a modified version of Uhlig's implementation [3, Ch. 4-5] is used.

(ii) An alternating minimization process is used to estimate both the set of rays intersecting the registered points on the monitor and to estimate the set of monitor poses ($\mathbf{R}_{\text{intr.}}, \mathbf{t}_{\text{intr.}}$) for the intrinsic calibration [3, Ch. 5].

3.2 Generic extrinsic camera calibration with ICP

The goal of the extrinsic calibration is to determine the rotation matrix $\mathbf{R}$ and the translation vector $\mathbf{t}$, that describe the camera location and orientation in the extrinsic calibration setup, hereafter called scene. Contrary to the intrinsic calibration, the extrinsic calibration must be performed every time the camera is moved.

The ArUco marker corners on the boards are the known calibration points $\mathcal{P}$ in the 3D scene. The surrounding cameras take corresponding pictures of the scene. The calibration points $\mathcal{P}$ are detected in the image with the OpenCV `detectMarkers()` function [2]. This results in the 2D corner coordinates $\mathbf{h}$ of each marker in the given picture that correspond to the known 3D calibration points $\mathcal{P}$. With this the `SolvePnP()` function from OpenCV [2] is used to get an initial guess of the rotation matrix $\mathbf{R}_0$ and the translation vector $\mathbf{t}_0$. They are used as starting values for the ICP (iterative closest point) algorithm.

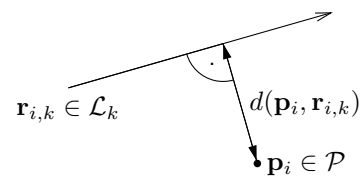


Figure 2: ICP method illustration for step k .

The ICP algorithm transforms the rays $\mathcal{L}$ to align with the known points $\mathcal{P}$. This process is described in [1] and uses a modified version of ICP originally introduced by Besl and McKay [19]. Furthermore it was adapted to align rays to points [20, Sec. 10.2.1]. This procedure is repeated for each camera individually. The procedure is explained in more detail in [1].

3.3 RMSE residuum in the ICP process

The extrinsic calibration accuracy is described by the RMSE of the minimum distance between a calibration target and the according calibrated camera ray for each ICP step:

$$\text{RMSE}_k(d(\mathbf{p}, \mathbf{r})) = \sqrt{\frac{1}{N_{\text{obs}}} \cdot \sum_i^{N_{\text{obs}}} [d(\mathbf{p}_i, \mathbf{r}_{i,k})^2]} \quad (2)$$

Here, N_{obs} is the number of observed points. $\mathbf{p}_i$ is the location of the i -th calibration point and $\mathbf{r}_{i,k}$ the according ray [1, 4, 14].

4 Improvements

4.1 Overexposure

Classical calibration approaches have mitigated this problem by using different patterns for calibration, namely checkerboard or ChArUco boards. The disadvantage of using such patterns is that they require a larger observed area without any occlusions [2, 9]. Our approach only requires individual ArUco markers to be fully visible. This was tested with camera poses that only see subsets of the markers and recordings with parts of the markers being obstructed [1, 4].

Furthermore we experienced in our setup, that glare due to spotlights and unsuitable camera positions negatively affected the detection quality. This is a common problem for various machine vision tasks, but especially for ArUco detection, as backed by the findings of Berral-Soler et al. [11]. This issue was addressed in the experiment setup by using diffuse lighting over a large surface area [13, Sec. 2.4.2]. We experienced, that for various exposure levels, the detected location of the marker corners slightly shifted towards the center of the markers with increasing exposure time.

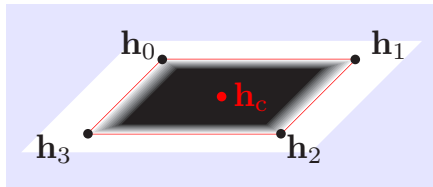


Figure 3: Schematic illustration of the blurred ArUco marker edge. The marker's internal structures are omitted. The red border is the real border, the shading illustrates the perceived brightness due to blur and defocus.

This can be explained by the blur recorded along the marker's boundaries (see Figure 3). To counteract this effect, we used the mean of the four corner coordinates as reference points, as we assume that the effect of blur is nearly equal for all four corners and the perceived shift of the four corners would cancel out.

$$\mathbf{h}_c = \frac{1}{4} \sum_{j=0}^3 \mathbf{h}_j \quad (3)$$

As opposed to the implementation of Agnus et al. [9] this proposed method does not require adjacent markers in a certain geometric position.

4.2 Subpixel mixture of rays

To utilize the **subpixel accurate corner detection** [2] to its full extent, we propose a method of mixing the four adjacent rays $\mathbf{r}_j = \mathbf{r}_0 + \lambda \mathbf{r}_a$ with $\mathbf{r}_0$ being a point on the ray and $\mathbf{r}_a$ its direction using bilinear interpolation.

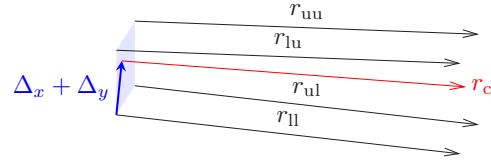


Figure 4: Illustration of the mixture of four adjacent rays.

The next lower or upper adjacent points are denoted with l and u respectively and the deltas denote the relative distance to the lower adjacent point $\mathbf{h}_{ll}$ [21, Ch. 2.3]. The combination resembles a weighted mean of the four rays. This procedure was performed both for the ray's origin $\mathbf{r}_0$ as well as its direction vector $\mathbf{r}_a$.

$$\begin{aligned} \mathbf{r}_{0,c} = & \mathbf{r}_{0,ll}(1 - \Delta x)(1 - \Delta y) \\ & \mathbf{r}_{0,ul}(\Delta x)(1 - \Delta y) + \\ & \mathbf{r}_{0,lu}(1 - \Delta x)(\Delta y) + \\ & \mathbf{r}_{0,uu}(\Delta x)(\Delta y) \end{aligned} \quad (4)$$

$$\begin{aligned} \mathbf{r}_{a,c} = & \mathbf{r}_{a,ll}(1 - \Delta x)(1 - \Delta y) + \\ & \mathbf{r}_{a,ul}(\Delta x)(1 - \Delta y) + \\ & \mathbf{r}_{a,lu}(1 - \Delta x)(\Delta y) + \\ & \mathbf{r}_{a,uu}(\Delta x)(\Delta y) \end{aligned} \quad (5)$$

Due to the rays being normalized by Uhlig's implementation [3, Sec. 5.2], the support vectors $\mathbf{r}_0$ are located near the optical center of the camera (where the rays are most dense). Therefore a weighted mean will find a suitable combined support vector $\mathbf{r}_{0,c}$. This method relies on the assumption, that the optical system of the camera is similar for adjacent pixels. Discontinuities between these pixels cannot be detected or treated. However, for regular camera geometries, this assumption is reasonable [3, Sec. 4.2].

5 Results

5.1 Overexposure

In Figure 5, we can observe the perceived position shift of the marker's corners viewed from an exemplary camera, based on its orientation in the scene. Each color represents one corner of a ArUco marker (top left, top right, bottom right, bottom left). We set the perceived position of the marker at the shortest exposure time as reference for a relative measurement.

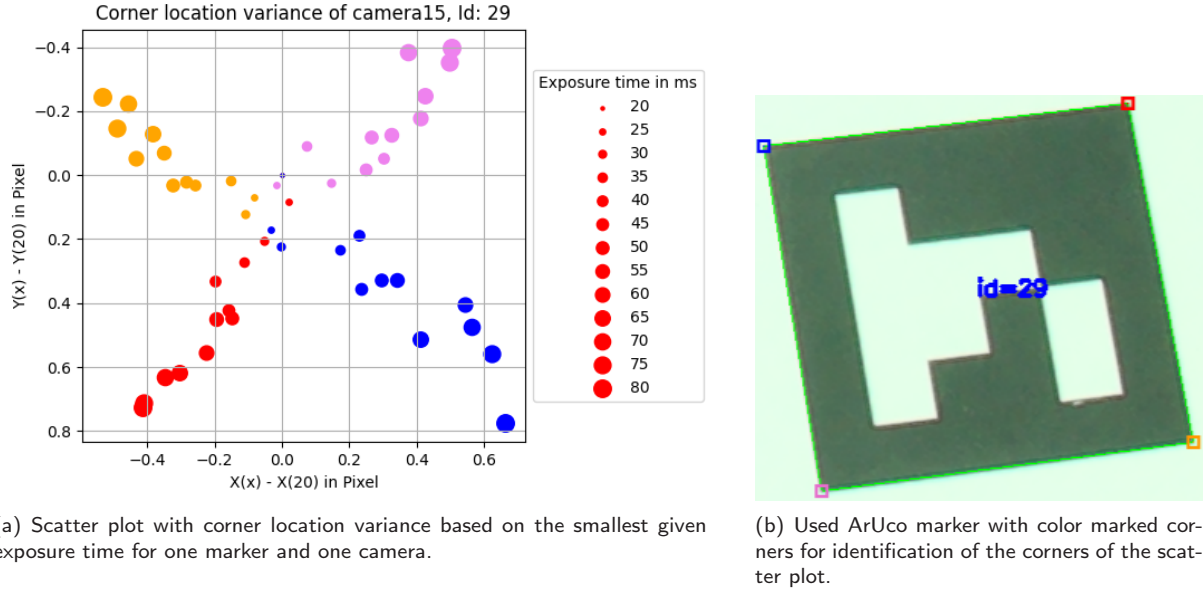


Figure 5: Corner location of an ArUco marker for different exposure times.

We can observe that each marker moves further away from its origin as exposure time increases. This hints to the fact that the perceived marker corners are detected more towards the center of the marker, due to the bright surroundings, introducing light to the nearby pixels. The observed effect of shift towards the marker center is the same as described by Agnus et al. [9] and Cledat et al. [7, Fig. 1].

The extrinsic calibration, using the corners of the marker as calibration points, results in a higher RMSE for higher exposure times. Especially if only one ArUco board is observed, the effect is the most noticeable.

5.2 Subpixel mixture of rays

To examine the influences of both measures (overexposure compensation by averaging of corners and usage of mixed rays) we present Figure 6. Here, the residual RMSE of the ICP process is plotted on the ordinate axis over various exposure times t_{exp} (abscissa).

The experiment was repeated for multiple individual cameras (columns), each individually intrinsically calibrated, and with the usage of the marker's corners directly (top row), as well as with the proposed weighted mean (bottom row). The graphs include both the values for utilizing full pixels and rays (blue) and the mixture of adjacent pixels (orange).

For the usage of the corners we can observe the rising RMSE associated with rising exposure times and thus most probably the effects of overexposure as described above. By utilizing the marker's centers, we can no longer observe this relation, as conjectured above.

Interestingly, the subpixel method does not result in any significant increase in calibration quality.

Furthermore, we can observe that the span of RMSE values is significantly smaller for the centers case as opposed to the usage of the marker corners. This is probably related to the process of averaging multiple observations of markers, thus reducing the measurement error.

Unfortunately the physical setup of the cameras does not allow for an accurate measurement of the optical centers of the cameras and thus the results cannot be verified exactly. However we measured the camera positions in the physical setup and considered them to be plausible.

6 Conclusion

In summary, we quantitatively evaluated the influence of overexposure on the process of generic extrinsic camera calibration.

Based on these findings, we improved the ICP method's results by utilizing the marker's centers and proposed a novel method for subpixel accurate ICP-based generic extrinsic calibration.

Our results indicate a strong dependency of the perceived marker corner location with varying exposure times. However, we found that the proposed approach of using the marker's center diminishes this relation and yields better RMSE results. Thus more robustness and accuracy for different lighting conditions can be expected for future experiments.

The usage of subpixels and mixed adjacent rays did not result in a significant decrease of the RMSE values. This might be an effect of image distortions as they are not corrected by the generic calibration. However for small distances in the image, the effect

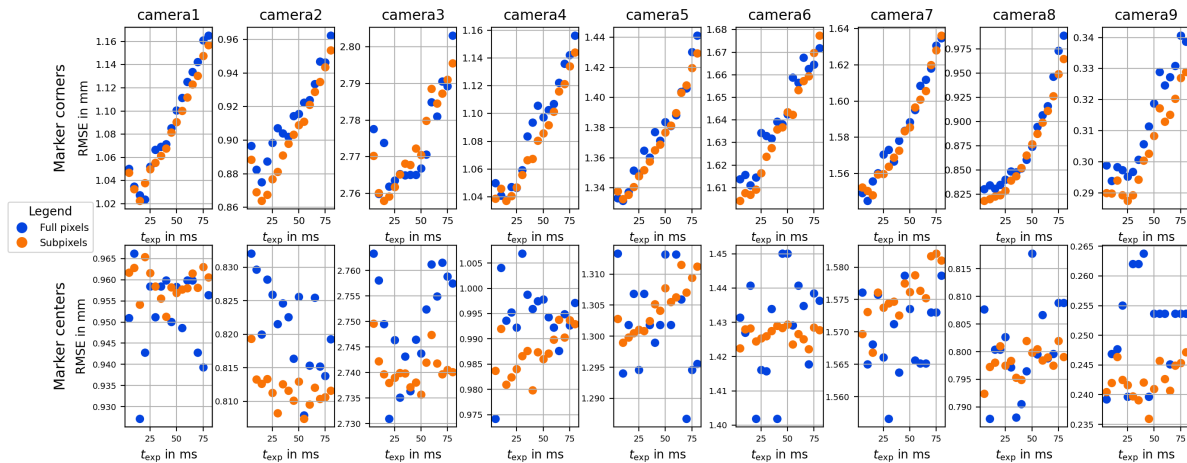


Figure 6: Resulting extrinsic RMSE for each camera for all combinations of calibration methods. Note the different scaling of the ordinate values.

is expected to be minor.

However there are still issues that future research should address: The extrinsic calibration result cannot be accurately verified with the current physical setup. Additionally, the current measure for accuracy, the RMSE residual distance between rays and calibration points might not be ideal as it allows for false local minima. Furthermore there might be other physical or optical effects influencing the accuracy of this process, such as image noise, chromatic aberration or systematic errors during the intrinsic calibration.

Moreover, the approach of overexposure-compensated calibration could be realized with different marker geometries, as proposed by Cledat et al. [7]. Especially the combination with generic camera calibration could result in more precise camera calibration.

Acknowledgments

- The authors would like to thank the Ministry of Science, Research and Arts of the Federal State of Baden-Württemberg for the financial support of the projects within the Innovation-Campus Future Mobility.
- The authors would like to thank the mechanical workshops at the IIIT and IRS for the manufacturing and technical support.

References

- [1] M. Zürn et al. “Robotic Wiring Harness Bin Picking Solution using a Deep-Learning-based Spline Prediction and a Multi-Stereo Camera Setup”. In: *SCAP2024* (2024).
- [2] G. Bradski. “The OpenCV Library”. In: *Dr. Dobb’s Journal of Software Tools* (2000).
- [3] David Uhlig. “Light Field Imaging for Deflectometry”. PhD thesis. Karlsruher Institut für Technologie (KIT), 2023. DOI: 10.5445/IR/1000156025.
- [4] Carsten Schmerbeck and Michael Heizmann. *Anwendung generischer Kamerakalibrierung für optische Messsysteme in der Fertigungstechnik*. German. 126. Jahrestagung der Deutschen Gesellschaft für angewandte Optik (DGaO 2025), Stuttgart, 11.–14. June 2025. 2025.
- [5] Thomas Schöps et al. *Why Having 10,000 Parameters in Your Camera Model is Better Than Twelve*. 2019. DOI: 10.48550/ARXIV.1912.02908.
- [6] M. D. Grossberg and S. K. Nayar. “A general imaging model and a method for finding its parameters”. In: *Proceedings 8. ICCV 2001*. Vol. 2. ICCV-01. IEEE Comput. Soc, 2001, pp. 108–115. DOI: 10.1109/iccv.2001.937611.
- [7] E. Cledat, M. Rufener, and D.A. Cucci. “Compensating over- and underexposure in optical target pose determination”. In: *Pattern Recognition* 116 (2021), p. 107930. ISSN: 0031-3203. DOI: 10.1016/j.patcog.2021.107930. URL: <https://www.sciencedirect.com/science/article/pii/S0031320321001175>.
- [8] Seyed Jalaeddin Mousavirad et al. “Addressing Contextual Factors in ArUco Marker-Based Distance Estimation: A Machine Learning Approach”. In: *2025 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*. IEEE, May 2025, pp. 1–

6. DOI: 10.1109/i2mtc62753.2025.11079149.
- [9] Vincent Agnus, Stéphane Nicolau, and Luc Soler. "Illumination independent marker tracking using cross-ratio invariance". In: *Proceedings of the 20th ACM Symposium on Virtual Reality Software and Technology*. VRST '14. ACM, Nov. 2014, pp. 69–72. DOI: 10.1145/2671015.2671119.
- [10] Víctor Manuel Mondéjar-Guerra et al. "Classification of Fiducial Markers in Challenging Conditions with SVM". In: *Pattern Recognition and Image Analysis*. Ed. by Luís A. Alexandre, José Salvador Sánchez, and João M. F. Rodrigues. Cham: Springer International Publishing, 2017, pp. 344–352. ISBN: 978-3-319-58838-4.
- [11] Rafael Berral-Soler et al. "DeepArUco++: Improved detection of square fiducial markers in challenging lighting conditions". In: *Image and Vision Computing* 152 (Dec. 2024), p. 105313. ISSN: 0262-8856. DOI: 10.1016/j.imavis.2024.105313.
- [12] Raspberry Pi. *RPI Documentation - Camera*. 2024. URL: <https://www.raspberrypi.com/documentation/accessories/camera.html> (visited on 03/27/2024).
- [13] Carsten Schmerbeck and Michael Heizmann. "VAMOS-Cam: Versatile Affordable Modular Stackable Camera System". 2025. DOI: 10.17632/75x3v4jyr2.1. in preparation.
- [14] Carsten Schmerbeck, Hesheng Wang, and Michael Heizmann. "Optimal aperture settings for generic camera calibration". In: *tm - Technisches Messen* 92.s1 (2025), pp. 93–98. DOI: 10.1515/teme-2025-0037.
- [15] G.AL® C250 – Aluminium Präzisionsplatte. GLEICH Aluminium GmbH, 2023. URL: https://www.gleich-aluminium-shop.de/module/mysydeshop/webforms/download_file.php?file=23 (visited on 10/16/2025).
- [16] *Messtechnik-Katalog*. Helios Preisser, 2023. URL: <https://metrology.mahr.com/fileadmin/catalogs/OnlineCatalogs/Helios-Preisser/index.php?catalog=HP-Metrology-DE&lang=de> (visited on 10/16/2025).
- [17] DIN Deutsches Institut für Normung e.V. *DIN 876-2:2014-12 - Prüfplatten; Prüfplatten aus Gußeisen; Anforderungen, Prüfung*. German Institute for Standardization (DIN). Berlin, 1984.
- [18] H.-J. Warnecke and W. Dutschke. *Fertigungsmeßtechnik. Handbuch für Industrie und Wissenschaft*. Ed. by H.-P. Grode. Vol. 1. Springer, 1984, p. 814.
- [19] P. J. Besl and Neil D. McKay. "A method for registration of 3-D shapes". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 14.2 (1992), pp. 239–256. DOI: 10.1109/34.121791.
- [20] P. J. Schneider and D. H. Eberly. *Geometric Tools for Computer Graphics*. Elsevier Science, 2003.
- [21] Fernando Puente León. *Messtechnik. Grundlagen, Methoden und Anwendungen*. 11th ed. Springer Vieweg, 2019, p. 481.