Assessment and Improvement of Distance Measurement Accuracy for Time-of-Flight Cameras

Valens Frangez, David Salido-Monzú, and Andreas Wieser

***Abstract*— Time-of-flight depth cameras are interesting sensors for contact-less 3-D metrology because they combine mechanical robustness with independence of ambient lighting conditions. Their actual performance depends on many factors and is hard to predict from data sheets. In this study, we investigate the deviations of the distance measurements of a high-end phase-based depth camera. We focus on the impact of: 1) self-warming and external temperature; 2) on range noise as a function of distance and acquisition time; and 3) on distance-dependent biases. We present the dedicated experimental setups comprising a climate chamber, a calibration bench with a reference interferometer, and a laser tracker that provides controlled conditions and ground-truth data. These setups allow investigating the absolute accuracy and mitigating repeatable distance biases by adapting the measurement model based on experimental data. For demonstration, we apply the investigation to two state-of-the-art industrial depth cameras of the same brand and type (Helios Lucid), showing significantly different response to external temperature but similar distance-dependent biases. We adapt the measurement model of one of the cameras for distance-dependent interpixel biases and demonstrate that the resulting parameters reduce also the distance biases of the other camera by about 80% to less than 1 mm at ranges of up to 1 m. This indicates the potential for batch error compensation. This article contributes to better understanding distance deviations of depth cameras and to improving the accuracy of such cameras.**

***Index Terms*— Accuracy assessment, depth camera, distance errors, error compensation, metrology, time-of-flight (ToF).**

I. Introduction

DEPTH camera technology has been commercially available for more than a decade and its use in various fields, such as manufacturing [1], mobile robotics [2], object detection [3], medicine [4], and scene understanding [5], has become an important part of many applications. The 3-D geometry acquisition within these fields usually refers to measurement ranges of a few meters with typically required accuracies of a few millimeters up to several centimeters. Various depth camera measurement principles, such as stereo triangulation, structured light, or time-of-flight (ToF) approaches, can be used for capturing the scene [6]. In particular, ToF-based cameras have gained attention in recent years, mainly due to their relatively low cost, long measurement ranges, high acquisition rates, and independence of the external illumination conditions. Since they became available, these cameras have represented a promising 3-D vision measurement solution for contactless metrology applications. Recent technological advances at the detector level have pushed the maximum range and speed of direct delay measurements (see, e.g., Sony IMX459[1] or SwissSPAD2 [7]) and improved the short-range precision of phase-based measurements such as the ones used in this article (see Sony DepthSense[2]). Given the typical measurement ranges up to a few meters and required accuracies at the millimeter level, the latter is driving a new generation of ToF cameras (see, e.g., Helios Lucid [8]) with increased potential in new application domains such as, e.g., geodetic and industrial metrology [9], [10].

In these contexts, the applications require a certain level of measurement stability over time, as well as sufficiently high precision and often high absolute accuracy. Depth camera accuracy is routinely improved by model adaptation, which can be carried out by the user [11] or by the manufacturer. For practical reasons, the latter process is likely done in bulk and not per device, often leaving room for further improvement as noted in [12] for the Helios Lucid camera.

The distance accuracy limitations of ToF depth cameras contributing to the overall measurement uncertainty can be grouped depending on the origin of the uncertainty into those driven by the sensor itself or its specific configuration, the measurement environment, and the geometry and radiometric properties of the measured scene [13], [14]. Lachat *et al.* [15] studied various measurement errors in ToF camera data, including distance-related systematic errors for which a simple error compensation approach was proposed using a total station. The ground-truth data were expressed with respect to the first acquired dataset, meaning that the accuracy was only quantified in a relative and not in an absolute manner. Distance errors as a function of the measured signal intensity were studied by Lindner *et al.* [16], who also proposed a correction function based on the intensity. An investigation on temperature-related distance errors, as well as distance-related errors, was carried out in [17] using a calibration setup and focusing on relative accuracy performance. Error compensation in depth cameras making use of additional

Manuscript received February 5, 2022; revised March 22, 2022; accepted March 29, 2022. Date of publication April 18, 2022; date of current version May 3, 2022. This work was supported by the Swiss National Science Foundation (SNSF) as part of the NCCR Digital Fabrication under Grant #51NF40-182887. The Associate Editor coordinating the review process was Dr. Chao Wang. *(Corresponding author: Valens Frangez.)*

The authors are with the Institute of Geodesy and Photogrammetry (IGP), Stefano-Franscini-Platz 5, 8093 Zurich, Switzerland (e-mail: frangezv@ethz.ch).

Digital Object Identifier 10.1109/TIM.2022.3167792

[1] www.sony-semicon.co.jp/e/news/2021/2021090601.html, accessed March 16, 2022.

[2] www.sony-depthsensing.com, accessed January 3, 2022.

measurements of the same scene acquired with RGB sensors has also been addressed [18], [19]. Frangez *et al.* [20] addressed the influences of the warm-up time and analyzed the dependence of systematic errors with distance, again in a relative manner expressed with respect to the first measured dataset. The warm-up time, the intensity-related errors, the signal-to-noise ratio (SNR), and relative distance-related errors are studied and compared for different depth cameras in [21]. To the best of the authors' knowledge, there is not yet any available work where an investigation on general performance assessment was carried out, which would quantify depth camera accuracy expressed in an absolute manner with respect to an accurate ground-truth reference.

Analyzing and understanding dominant error sources and defining adequate device measurement model adaptation procedures is one of the main tasks in distance metrology instrumentation, where the error compensation of electro-optical distance measurement (EDM) devices, such as, e.g., total stations [22] is probably the most established example. Distance-related errors, such as the impact of variations of reflectivity and systematic deviations on reflectorless measurements using total stations, along with the assessment of the modeling potential of systematic errors were studied in [14]. Lichti and Lampard [23] proposed a self-calibration approach for total stations. Similarly, distance effects, such as cyclic deviations and scale variations in terrestrial laser scanners, were investigated in [24]. These works highlight the opportunities for achieving significant performance improvement on commercial EDM instruments by carrying out measurements and procedures for error compensation.

The main contributions of the presented study are twofold.

1) We propose an approach for performance assessment of ToF cameras by considering sensor and environment-related relevant sources of uncertainty. Ground-truth data are realized using metrology setups, instrumentation, and procedures and are expressed in relation to the internal camera coordinate system (CS), allowing for reproducible and absolute accuracy quantification. This approach is then applied to the state-the-art industrial ToF depth cameras Helios Lucid.
2) We use the analysis outcomes to derive a measurement model adaptation to correct for repeatable systematics, considering distance and interpixel-related errors. The model is demonstrated to be transferable between two cameras of the same model providing a significant accuracy improvement and signaling potential for the generalization of our approach.

A short overview on the phase-based ToF measurement principle of depth cameras and sources of uncertainty is given in Section II. The measurement setups and data processing pipelines are described in Section III. In Section IV, the results are presented and discussed, and an approach for error compensation is presented. Conclusions are given in Section V.

II. Measurement Principle and Sources of Uncertainty

The ranging process of ToF depth cameras relies on measuring the accumulated phase of an intensity-modulated optical signal illuminating the scene and detection on each pixel [6], [25]. The scene illumination is typically realized by several near-infrared LEDs collocated with the camera sensor and driven by a radio frequency envelope. The reflected illumination signal is through optics focused on the imaging chip, where the signal is demodulated at the device level for each of the pixels within the chip. This demodulation is carried out by computing the correlation between the detected signal and the internal reference clocking the modulation of the LEDs [26]. The accumulated phase is derived from the correlation samples acquired over a certain interval, i.e., the integration time, and, given the modulation frequency and the signal propagation speed, the distance $\hat{d}_k$ for each pixel k can be calculated therefrom. The cyclic nature of the measurements makes them inherently ambiguous beyond the modulation wavelength, which is typically overcome in practice by combining simultaneous observations at more than one modulation wavelength.

The measured distance $\hat{d}_k$ is then used along with the geometric model of the camera for mapping in 3-D space, computing the Cartesian coordinates of the measured point as

$$\begin{pmatrix} X_k \\ Y_k \\ Z_k \end{pmatrix} = \begin{pmatrix} x_k \\ y_k \\ 0 \end{pmatrix} + \begin{pmatrix} -x_k \\ -y_k \\ c \end{pmatrix} \cdot \frac{\hat{d}_k}{\sqrt{x_k^2 + y_k^2 + c^2}} \tag{1}$$

where x_k and y_k are sensor coordinates and c is the camera constant. These three parameters are a part of the intrinsic camera matrix resulting from a standard camera calibration process [27] using, e.g., a checkerboard target. The conversion using the perspective projections using (1) is based on a simple pinhole camera model, which is a sufficient approximation in this case [13], [17]. Once computed for all the camera pixels, the group of measured 3-D coordinates is provided to the user as an individual acquisition. Each acquisition results in a depth image ($[x_k \; y_k \; \hat{d}_k]_{k\times 3}$), which can be using (1) transformed into the 3-D point cloud ($[X_k \; Y_k \; Z_k]_{k\times 3}$) point-by-point. Focusing only on the ranging performance—as opposed to the full 3-D performance including also angular uncertainties due to, e.g., image distortion or irregular pixel distribution—Fig. 1 shows a comprehensive list of plausible sources of uncertainty for ToF depth cameras. We have identified the sources of uncertainty considering the underlying measurement principle and the available literature on depth camera error compensation and analysis [15]–[21], [28], [29]. The listed uncertainties, further specified in the following paragraphs, are classified depending on whether the impacting phenomenon or parameter is defined by: 1) the sensor itself; 2) its specific measurement configuration; 3) the measurement environment; or 4) the acquired scene.

1) Sensor: The SNR of noncoherent optical detection systems, such as ToF cameras, is mainly defined by the relationship between a quasi-constant additive noise contribution

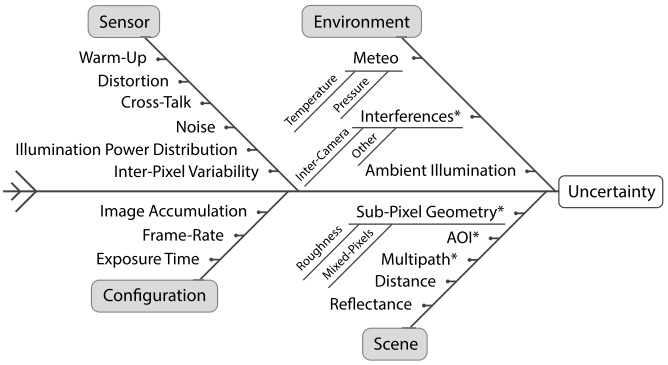


Fig. 1. Ishikawa diagram with the identified sources of uncertainty for the distance measurements of ToF depth cameras (* indicates phenomena that are not directly or indirectly addressed in this study).

at the photodetection/amplification stages and the detected optical power [21]. The short-term precision due to random contributions thus depends on the illumination signal power on the area of interest and the radiometric and geometric properties of this area [16], [20], [28]. In addition, the illumination signal power distribution along with interpixel variations of the sensitivity and time response of the sensor may introduce pixel-dependent precision variations and systematics, respectively [20]. Distance-dependent systematics are also typically present in phase-based measurement systems, introducing trends and cyclic errors on the observed ranges [15]–[19], [21]. These are generally produced by internal electrical and optical interferences within the sensor due to crosstalk of the illumination signals onto the received ones. These systematics are specifically aggravated in ToF cameras due to distortion of the illumination signal, which is typically processed assuming pure sinusoidal modulations when actually also composed of significantly high-order harmonics.

2) Configuration: Modern ToF cameras usually provide selectable or automatically configured frame rates. Different frame rates set the exposure and integration time used for the individual depth image acquisitions, balancing noise reduction and dynamic performance while avoiding saturation of the sensor due to ambient illumination [15]–[17], [29]. In addition to signal integration within individual acquisitions, consecutive depth images can be averaged in post-processing given a sufficiently static scene to further increase precision.

3) Environment: The density of the air and its spatial variation affect the speed of signal propagation along the line-of-sight (LOS) [30] and the curvature of the LOS [31]. However, due to the short distances, the curvature can be neglected for ToF measurements and the signal propagation speed can be calculated for constant, representative meteorological conditions. The impact of the measurement environment is thus mostly determined by the temperature, pressure, and illumination conditions. Temperature changes, both external and caused by self-warming of the instrument, induce mechanical and electrical variations in the sensor, which may produce significant temperature-dependent systematics [15], [17], [20], [21]. The impact of the pressure is typically low and can be considered negligible, as, e.g., it will not exceed 0.15 mm over 10 m with maximum expectable change on a fixed location of 50 hPa. The impact of ambient illumination on the other hand is twofold. Excessive background light may saturate the sensor, reducing its dynamic range and therefore preventing distance measurement [15]. This is usually prevented in practice by using lower exposure times on outdoor measurements at the cost of reduced precision and reduced range. In addition, ambient illumination from oscillating sources such as fluorescent tubes may interfere with the illumination signals when a multiplicity relationship with the modulation frequencies exists. Modern ambient illumination systems, however, are increasingly less problematic due to their spectral distribution and continuous-power operation. An additional and occasionally critical impact of the environment arises from external interferences on the sensor, which may be significant in environments with high electromagnetic radiation such as, e.g., industrial settings with large electric motors. The simultaneous use of several cameras of the same type may also introduce errors due to optical interference between the illumination signals. This is solved in modern cameras by enabling several modulation frequency configurations to enable multisensor measurements without intercamera coupling.

4) Scene: Since depth cameras carry out measurements using the optical signals backscattered by natural targets, some uncertainties are dependent on the interaction between sensor features and geometric and radiometric properties of the measured scene [15]–[17], [28]. Along with the illumination signal, the combination of distance, angle of incidence, and reflectance of the surface areas covered by the individual pixel determine the collected optical power and thus the precision. The total collected power together with the return phase of the illumination signal also impacts the systematics due to signal distortion and internal interferences. Equivalently to interferences from another camera, the geometry of the environment may also cause multiple reflections—typically referred to as multipath interferences—biasing the measured distance in, e.g., corners and cluttered scenarios. Finally, the inherent spatial averaging of the geometry within each pixel (i.e., subpixel geometry) in relation to the camera resolution induces biases with respect to the assumed true distance when the covered area contains significant distance variability due to roughness or object edges, resulting in the so-called mixed-pixel measurements [16], [17].

Our focus in this work is on analyzing and quantifying the expected accuracy of a state-of-the-art ToF camera independently of the measurement scenario and investigating mitigation strategies that can be extrapolated to other systems. We base our analysis on experimental investigations of the obtained ranging performance by direct comparison with sufficiently accurate ground-truth measurements in controlled scenarios. We target the identification of random and systematic uncertainty components and analyze the reduction of both by signal integration and measurement model adaptation, respectively. The tests realized here, therefore, address the joint impact of sensor-, configuration-, and environment-dependent effects. We exclude scene-related parameters, whose analysis is largely application-dependent and whose mitigation can only be addressed at the point cloud level—as opposed to the individual measurement level.

TABLE I
SELECTED SPECIFICATIONS OF THE HELIOS LUCID DEPTH CAMERA [8] (ACCURACY IS DEFINED AS AVERAGE DEVIATIONS FROM GROUND TRUTH AND PRECISION IS DEFINED AS REPEATABILITY)

Field of View	59° × 45°
Resolution	640 × 480 pixels; 0.3 MP
Range	0.3 m to 6.0 m
Accuracy	±5 mm for up to 1.5 m
Precision	0.5 mm (<0.5 m); 2.2 mm (<1.5 m)
Illumination	4 VCSEL laser diodes @ 850 nm
Modulation Freq.	70 MHz and 100 MHz

III. EXPERIMENTAL SETUP AND METHOD

A. Depth Camera

The proposed approach of performance assessment was demonstrated using a state-of-the-art industrial phase-based ToF depth camera type (Helios Lucid [8]), which operates on the measurement principle described in Section II. The most relevant specifications of this camera type are summarized in Table I. Two different cameras of that same model were used across the presented study to compare their performance, analyze the intercamera agreement, and assess the generalization potential of the derived conclusions.

The camera provides two different acquisition modes that define a maximum working range of 1.5 and 6 m, where the latter combines the two modulation frequencies given in the table to extend the unambiguous measurement range. All of the measurements for the presented study are carried out using the 6-m measurement mode to maximize the range covered by the assessment. The camera provides selectable frame rates, i.e., how many single frames are averaged for the individual output depth and intensity images, representing the geometry of the observed scene and the power backscattered from the respective surface, respectively. Within our analysis, each single depth image used in the evaluation is derived by default from averaging over 30 continuous frames, i.e., using the 15-frames/s frame rate acquired over 2 s. The exposure time can also be configured to 250 or 1000 μs, of which the former was selected to avoid saturation on highly reflecting surfaces at short ranges. These parameters were selected as a tradeoff between noise absorption and experiment duration, and specific tests on noise reduction as a function of image accumulation were included in the assessment. Both cameras are operated via the Ethernet cable connected to a laptop and controlled by a Python script.

B. Measurement Setup

The data acquisition was carried out in two different metrological facilities, namely, a climate chamber and a calibration bench for distance metrology. The climate chamber was used to collect the data necessary to investigate temperature-related effects, namely, the impact of warm-up and external temperature on the measured distances. The calibration bench, on the other hand, provided the basic mechanical platform and reference system to study all the distance-related effects and was additionally used for the static data acquisitions required for precision analysis. The two setups are described in detail next.

1) Climate Chamber: It allows for varying temperature and relative humidity in a controlled manner from −30 °C to 50 °C and from 40% to 80%, respectively. A camera was set up on a metal pole separated 60 cm from a metal plate used as a measurement target. The plate was coated with an antireflection spray MR 2000 Anti-Reflex L for 3-D metrology applications to guarantee diffuse and homogeneous reflectance. We estimated the depth errors due to temperature-dependent deformations of the plate and supporting pole, which are estimated to be smaller than 0.1 mm, considering the used alloys and the design. We considered this to be negligible for the present investigation.

2) Calibration Bench: It consists of a linear horizontal bench hosting a motorized trolley that can move along the bench. The position of the trolley relative to its starting point on each run is controlled by a computer and a Doppler interferometer. The accuracy of the interferometer measurements and the position of the trolley is better than 30 μm. A similar measurement setup using the bench was used in related experimental studies in, e.g., [14] and [32]. The bench is located in a room with stable temperature and humidity. In particular, the temperature variability is less than 0.1 °C.

The depth camera was mounted on the moving trolley and an 841 mm × 1189 mm (A0) metal plate used as a target was fixed on one end of the bench. The planarity of the plate was assessed by acquiring points on its surface using a laser tracker (LT) (Leica AT960) and a contact probe (Leica T-Probe). The results indicated root-mean-square (RMS) deviations of the plate of less than 0.1 mm from a perfect plane. The trolley experiences small tilt variations when moving along the bench, with vertical variations bounded to 0.31 mrad.

The calibration bench provides accurate ground truth relative to the first position of the trolley on each run. The goal of the analysis is to assess the absolute accuracy of the distance measurements. Therefore, the missing link is the establishment of an absolute distance between the target and the distance measurements of the camera for one of the target positions. Thus, to connect the relative measurements from the interferometer with the camera data, we used the LT and the contact probe to directly determine the position and orientation of the external coordinate system (ECS) of the camera with respect to the target plate, while the trolley was in its initial position. Finally, we related the internal coordinate system (ICS) of the camera to the ECS by determining the transformation matrix $\mathbf{M}_{\text{ECS}}^{\text{ICS}}$ using the hand–eye calibration method from [9] and [33]. Overall, this means that we are able to extract ground-truth distances with deviations of less than about 0.1 mm for all pixels of the ToF camera during the later analysis.

C. Data Processing

Once the distance $\hat{d}_k$ and its direction vector $\vec{d}_{k,1} = [X_k, Y_k, Z_k]$—computed as the LoS direction of the corresponding pixel—are recovered, we can compare those measured distances with the ground-truth data to obtain the

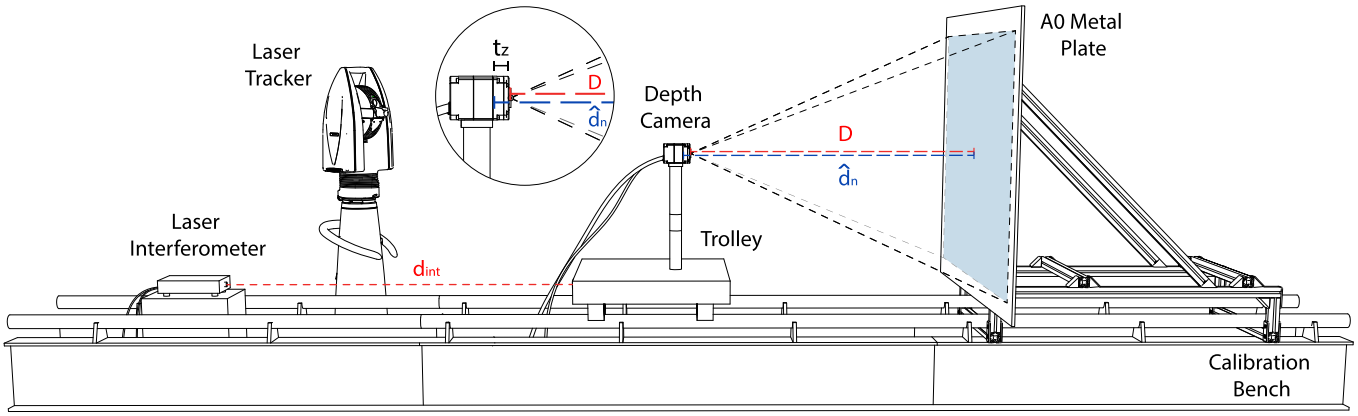


Fig. 2. Calibration bench measurement setup: d_{int} is the laser interferometer measurement, D is the offset between the plate and the ICS and measured using the LT, and $\hat{d}_n$ is the depth camera distance measurement for the central pixel n. The close-up on the depth camera shows the different parameters used for calculating the distance correction $d_{n,\text{dev}}$ [see (2)]. D and $\hat{d}_n$ slightly displaced with respect to the camera's front face are for visualization purposes.

difference values representing absolute errors. First, plane parameters A, B, C, and D are determined based on the LT measurements. The target plate is defined in the same CS as the ECS and, by knowing the $\mathbf{M}_{\text{ECS}}^{\text{ICS}}$, also in the ICS. In the next step, we calculate the intersection between the plane and the vector $\vec{d}_{k,1}$, resulting in an intersection point that represents the true measurement location and is given by the vector $\vec{d}_{k,2}$, which points in the exact same direction as $\vec{d}_{k,1}$. We then subtract the lengths of the two vectors and obtain their length difference, i.e., the deviation $d_{k,\text{dev}}$. This operation is done on a per-pixel basis. Since the trolley moves in steps denoted as j, the plane parameter D has to be constantly updated and a new plane has to be computed for each comparison of the depth camera dataset acquired on that respective location. For each step, the trolley displacement $d_{\text{int},j}$ is determined based on the laser interferometer measurements d_{int} between the current step j and the first step and is simply added to D. In a general form, the absolute distance deviation for the central pixel denoted here as n (see Fig. 2, the diagram of the measurement setup with all marked relevant parameters) can be expressed using

$$d_{n,\text{dev}} = \hat{d}_n - t_z - \left(D + d_{\text{int},j}\right) \tag{2}$$

where $\hat{d}_n$ is the measured distance for the central pixel and t_z is the Z-component of $\vec{t}$ of the transformation matrix $\mathbf{M}_{\text{ECS}}^{\text{ICS}}$. This represents a simplified definition for a case not considering the orientation of the plane with respect to the ECS, which is in fact accounted for in the line-plane intersection computation.

IV. Results

The processed results for all the experiments are presented and discussed in this section. First, the impact of temperature, both external and self-warming, is investigated along with the possibility of compensating temperature-dependent deviations using the internal readings from the camera. The measurement precision and its improvement by image averaging given a sufficiently static scene are then assessed. Finally, the deviations as a function of the measured distances are analyzed, together with the precision degradation with distance. The results are interpreted and used to assess the potential for compensation of systematic deviations.

A. Deviations With Temperature

To analyze the impact of self-warming, the camera was kept at a constant temperature of 25 °C in the climate chamber. The camera was acclimatized for several hours to the temperature within the chamber prior to the experiment taking place. Acquisitions triggered every 20 s were initiated immediately after the camera was turned on and were recorded over a period of 2 h. The internal temperature reading of the camera, with a resolution of 0.25 °C, was also recorded. The full experiment was repeated three times for the first camera and two for the second camera.

The results are shown in Fig. 3, depicting the estimated distance change to the target plane for a representative pixel along with its intensity and internal temperature readings. The representative pixel was arbitrarily selected to capture the general behavior of the majority of the pixels in the image with only negligible differences in their magnitudes and patterns. When not stated otherwise, the pixel is selected close to the center of the image and is the same for the five datasets shown in the figure. A low-pass-filtered version of the estimated distance is shown instead of the raw data to reduce noise, highlighting the much slower temperature-induced drifts. As seen in the figure, the warm-up period took approximately 40 min and the distance changed by around 2.4 and 1.3 mm for the first and second cameras, respectively, with respect to the median value of the first ten measurements before stabilizing. The distance change follows the same pattern of the acquired intensity. A relatively high correlation of the distance deviation with the internal temperature is also apparent from the results, although not sufficiently repeatable between cameras so as to justify developing a correction model for batch error compensation based on measurements with a single camera.

The impact of external temperature change was investigated by introducing controlled variations of the chamber temperature from 15 °C to 35 °C over 55 h. The temperature profile comprises two cycles reaching the same temperature values

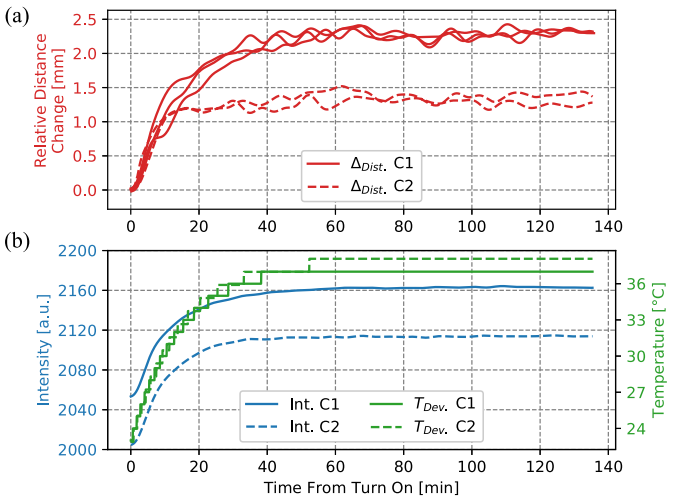


Fig. 3. (a) Distance changes for representative pixel caused by the warm-up process and (b) their corresponding intensity change and internal device temperatures of the two cameras, as a function of time from turn on. Three datasets are available for the first (C1) and second (C2) cameras.

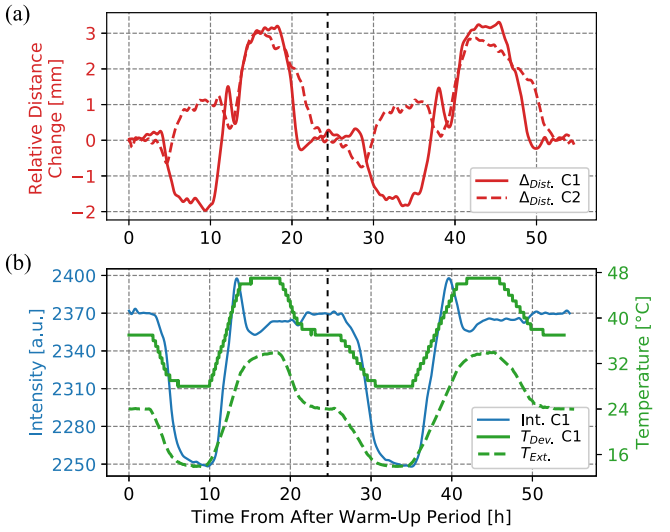


Fig. 4. (a) Distance changes for a representative pixel caused by external temperature variations and (b) their corresponding intensity change and internal device temperatures of the two cameras, as a function of time from after the warm-up period. C1 stands for the first camera and C2 stands for the second camera. A vertical dashed line approximately in the middle indicates where the second cycle of the temperature profile starts.

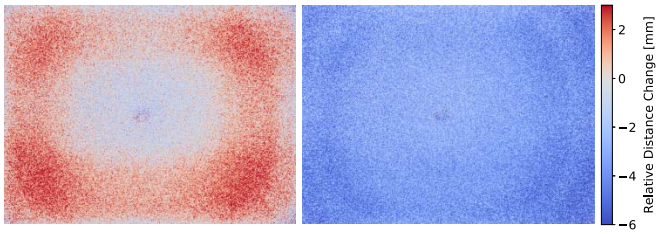


Fig. 5. Relative distance changes shown for the whole image for two instances, i.e., at the highest (left) and at the lowest (right) recorded internal temperature.

with a slightly different slope in the transition areas. Acquisitions were triggered every 20 s, together with the recording of the internal device temperature and the external temperature, which is as observed in our case lower by approximately 20 °C. The measured intensity and the distance change relative to the median of the first ten measurements, along with the internal temperature, are shown in Fig. 4. The experiment was carried out once for each of the cameras, resulting in two temperature cycles per camera. This allows assessing intracamera and intercamera repeatability of the impact of temperature.

As seen in the figure, there is high repeatability of the internal temperature readings on both cycles within the resolution of the sensor. In addition, the agreement between the internal temperature values for the two cameras is on the level of 1 °C (internal temperature for the second camera is not shown in the figure). When the internal temperature is stable within 2 °C for more than approximately 30 min (i.e., at 15 °C, 25 °C, and 35 °C), the intercamera repeatability between the two cycles is 0.5 mm or better (2σ). The variation pattern is, however, significantly different for the two cameras with deviations of up to 3 mm for the same stable temperature. The results suggest that compensating temperature-related errors down to the sub-millimeter level using the internal temperature readings is possible but requires deriving an independent error model per camera, which is likely prohibitive for most application cases.

Uncompensated, the variation of the external temperature-induced absolute distance changes of approximately 5 mm—about 1% of the absolute measured range of 60 cm. The observed distance variations cannot be justified by refractivity errors due to the changing propagation velocity of the optical signal, which would be well below 0.01 mm, considering the temperature span and the relatively short baseline of 60 cm. Thus, they are likely caused by mechanical changes within the camera due to thermal deformations. An independent test was conducted to guarantee that the observed magnitudes and patterns do not result from thermal deformations of the aluminum board used in the experiment. Different parts of the board were captured when the environment temperature was changed. We found only negligible deviations between these parts and thus presume internal camera errors as the origin of the observed deviations.

To visualize different interpixel behavior inducing a spatial distribution of the impact of temperature, Fig. 5 shows the estimated distance deviation over the whole depth image with respect to the beginning of the measurements for an instance of the highest and lowest internal device temperature. The inhomogeneity of the impact of temperature throughout the sensor is clearly visible on the different magnitude of the spatial patterns for both temperature extremes. The maximum and minimum values for a single pixel shown in Fig. 4 correspond to 3 and −2 mm, respectively; however, the values for other pixels within the depth image might be slightly different and appear within the range from 3 to −6 mm, as shown in Fig. 5.

Mitigating the impact of temperature to increase the absolute accuracy of the camera can therefore be primarily addressed at the user level, by guaranteeing sufficient warm-up time (approximately 40 min) before the actual measurements

and keeping the measurement environment stable within 2 °C. In this condition, sub-millimeter errors can be expected according to our results, which is well below the error magnitude due to other sources of uncertainty. Compensation of temperature-related errors to allow for larger temperature variability in the environment while maintaining sub-millimeter error levels is also possible but requires extensive effort and equipment to derive camera-specific error models.

B. Deviations With Distance

The analysis of distance-related deviations is based on measurements on the calibration bench setup described in Section III-B2. The trolley was moved along the bench over the complete measurement range of the camera from approximately 0.3 to approximately 6 m. The metal target plate was covered with highly diffusive paper films of four different colors (white and three shades of gray), using target reflectance as a proxy to assess the impact of different received power on the distance-related deviations. A relative reflectance of 100%, 95%, 82%, and 70%, for white, light, medium, and dark gray, respectively, was calculated experimentally from the average intensity values across the whole image measured at a fixed distance of 0.4 m. The position of the plate with respect to the camera was determined newly after each cover change. The trolley step size was set to 2 mm for the white cover, taking about 12 h for a complete measurement cycle, and 5 mm for the remaining reflectances, which was completed in about 5 h. The expected cyclic errors associated with the implemented modulation frequencies (75 and 100 MHz) are 2 and 1.5 m, respectively. With the chosen small step sizes potentially, present cyclic errors should be clearly visible in the results.

For the first camera, the complete measurement cycle was repeated five times on the white cover and once on the others, while only one cycle per cover was acquired with the second camera. The time span between the first and the last acquired dataset for the first camera was about four weeks. The field of view (FoV) of the camera covered the plate completely for up to approximately 1 m, with only the central part of the image being covered for the remaining range. To reduce the influence of the vibrations of the trolley after each displacement, a stabilization time of 5 s was allocated before triggering the depth camera measurements.

All the results shown in this section are given for an arbitrary but representative pixel and, unless otherwise noted, the derived conclusions hold also for all the other pixels. For a clearer interpretation of the results, all the distance deviations shown in this section are obtained by low-pass filtering the original distance series, aiming at reducing noise-related random deviations to isolate repeatable systematics in the data typically exhibiting much slower spatial frequencies. This was achieved using a low-pass Butterworth filter with the cutoff frequency of the filter selected experimentally as a tradeoff between maximizing noise absorption while preserving the variations of the clearly systematic components.

The extracted distance deviations for all five datasets on the white surface using the first camera are shown in Fig. 6.

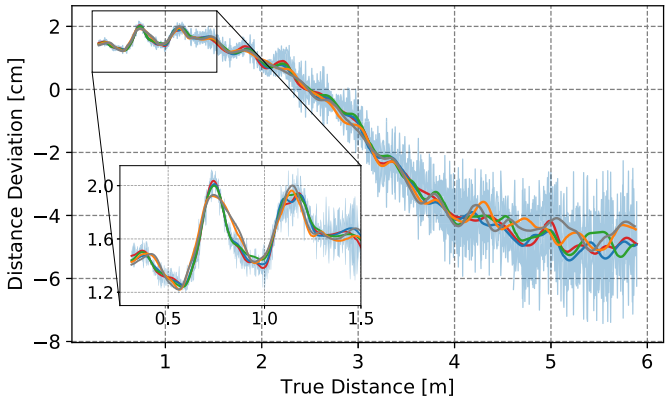


Fig. 6. Distance deviations as a function of distance for the five cycles acquired using the first camera on a white plate cover. The low-pass-filtered data for the different datasets are shown in solid lines, while the blue curve represents the original dataset for one of the five measurement cycles.

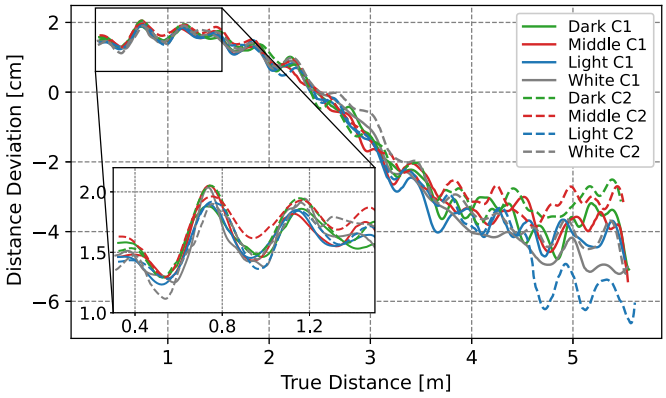


Fig. 7. Distance deviation as a function of distance for eight different datasets using both cameras (C1 and C2) surface covers with different reflectances.

The low-pass-filtered data for the different datasets are shown in solid lines, while the blue curve represents the original dataset for one of the five measurement cycles that still contains noise-related random deviations. The comparison of the low-pass-filtered datasets of all the measurement cycles shows repeatable patterns and trends as a function of distance, with no obvious relationship with the modulation wavelengths. The agreement between these datasets is better than 1 mm for distances of up to 1.5 m. Beyond this distance, the increase in noise introduces deviations on the low-pass-filtered data of a magnitude comparable to that of the smaller systematic patterns, hindering its fine evaluation. A general agreement for the coarser systematic trends is nevertheless clearly visible across the complete range. These results suggest room for accuracy improvement by measurement model adaptation. The following results aim at assessing whether these systematics are strongly dependent on the received power, suggesting the need to use the intensity data as an input for measurement model adaptation, and whether the repeatability holds across different cameras so that batch error compensation based on a single measurement campaign could be applied.

The results for the two cameras and all four different surface coverings are shown in Fig. 7. The deviation patterns of all datasets agree within 3 mm for up to approximately 2 m. For longer ranges, as in the previous results, noise dominates

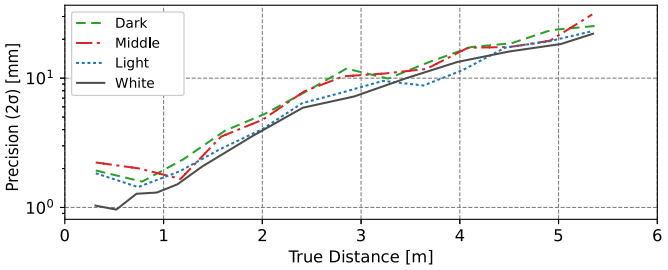


Fig. 8. Distance precision (2σ) as a function of distance for different surface reflectances.

over the smaller trends and the agreement is only visible on the coarser systematic trend. The results are independent of changes on the detected power induced by the different reflectance of the various surface covers. This may be due to an inherent immunity of the sensor to power changes or a built-in compensation of intensity-related systematics already applied by the manufacturer.

The results overall indicate a high potential for accuracy improvement via measurement model adaptation, which does not need to consider intensity information—thus likely highly immune to different target reflectances within the assessed range and partly also to varying incidence angles—and can be extrapolated between different models of the same sensor. This error compensation procedure is addressed in detail in Section IV-C.

C. Precision

The precision of the sensor and its dependence on distance, pixel, and measurement time was investigated using the previous results along with static acquisitions. The residuals from subtracting the low-pass-filtered results from the original data are dominated by the underlying range noise. In fact, they represent a slightly underestimated version of the original range noise due to the absorption of low-frequency noise components but can nevertheless be used as a proxy to investigate the dependence of range noise with distance. The measurement precision (2σ) per distance, computed from a window of 100 measurements, i.e., within a range of 20 cm, is shown in Fig. 8. Our estimated precision agrees closely with the precision (2σ) of 2.2 mm specified by the manufacturer for a range of 1.5 m. The measured intensity (not shown) accurately follows a quadratic decay with distance. As expected from this reduction of signal power, the standard deviation of range noise with distance shows also an approximately quadratic increase, reaching up to 3 cm when approaching the maximum working range of the camera. Range noise is also expected to be inversely proportional to target reflectance. The curves are generally following the expected order, with higher reflectance targets leading to lower range noise; however, the relatively small range of evaluated reflectances (ranging from 70% to 100%) does not allow for a better quantification.

Static measurements on the calibration bench placing the target with the white cover at approximately 70 cm from the sensor were used to assess the noise characteristics across the sensor and the efficiency of its reduction via image averaging. An acquisition was triggered every 4 s over 6 h,

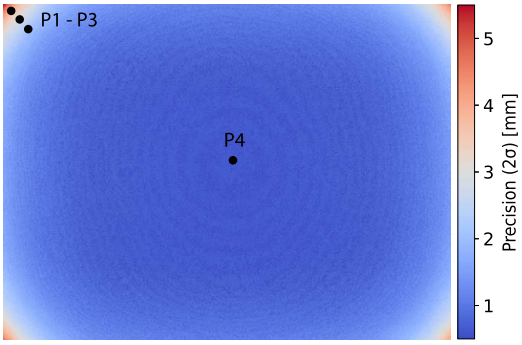


Fig. 9. Precision (2σ) of the measured distance per pixel computed from 6000 depth images acquired over 6 h.

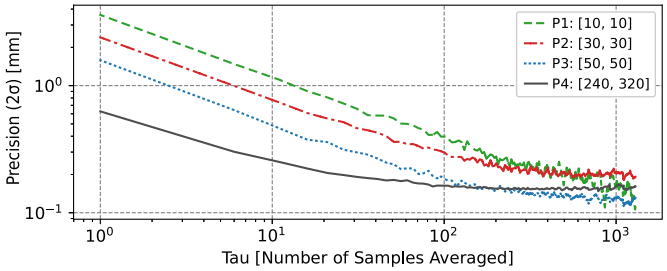


Fig. 10. Precision (2σ) as a function of number of averaged images computed for pixels $P1$, $P2$, $P3$, and $P4$ as marked in Fig. 9.

resulting in about 6000 depth images. The full acquisition cycle was repeated three times with the first camera and once with the second one.

Fig. 9 shows the estimated distance precision (2σ) per pixel computed empirically for one of the datasets of 6000 depth images. The estimated precision ranges from 1 to 5 mm, degrading radially across the image as expected considering the larger distance to the planar target and possibly an additional signal power reduction due to the smaller illumination and light collection efficiency of the sensor with increasing angle. A weak oscillating radial pattern is also apparent from the measurements, probably also induced by a nonhomogeneous spatial distribution of the illumination signal and the coupling efficiency of the detection optics.

When acquiring static scenes, if allowed by the application requirements, image averaging can be effectively used to reduce range noise. The efficiency of this precision improvement depends ultimately on the spectrum and correlation properties of the noise. We have investigated this experimentally by computing the distance precision (2σ) as a function of the number of contiguous individual measurements used for averaging. The resulting range noise for the four pixels marked in Fig. 9 is shown in Fig. 10. The results show that averaging up to approximately 100 depth images, corresponding to an acquisition time of 200 s, provides efficient noise reduction with approximately the square of the number of images. This indicates that the range noise is largely uncorrelated and of constant density across the spectrum except for very low frequencies. The averaging becomes significantly less efficient above 100 images, suggesting a lower limit on the achievable precision around a few hundreds of micrometers.

D. Measurement Model Adaptation

The high repeatability of the observed deviations with distance indicates the potential for significant accuracy improvement by error compensation. In addition, the large agreement on those systematics between both of the evaluated cameras suggests that measurement model adaptation may be practically suitable since the relatively high measurement effort could be leveraged by applying the resulting corrections to multiple sensors. Such model adaptation is addressed in this section by considering the distance and interpixel variations. To validate the transferability of the approach, the error compensation functions are derived from the distance-related measurements on the first camera and then applied to the second to assess the resulting corrected accuracy. We have restricted the model adaptation to the short range of the camera up to 1 m. This is, on one hand, the region where uncorrected systematics dominate over noise on the overall error budget, thus effectively improving the sensor performance and hence justifying the error compensation efforts. On the other hand, the complete FoV of the camera is only contained in the target plate within that range. A calibration setup able to cover the whole sensor across a longer distance would require an impractically large target that could not be combined with the highly accurate reference from the calibration bench and interferometer, thus defining a significantly higher error compensation effort likely unacceptable for any application. Nevertheless, the error compensation results presented in this section show potential for their generalization to distances between 1 and 6 m, as was investigated using the central image pixel, which always targeted the center of the plate; thus, its measurements are contained within each dataset acquired within 6 m. Anyhow, this part of the investigation will not be further discussed here and the focus will remain on the datasets acquired up to 1 m.

One full dataset of the measurements to the white cover was used to estimate a polynomial fit to the distance deviation as a function of the true distance independently for each pixel. A polynomial of degree 6 was fit for each pixel, as determined by the Akaike information criterion [34] to avoid data overfitting. The fitting quality was evaluated by computing the average coefficient of determination for all of the pixels for the distances covered by the model, yielding an average value of 0.99, thus indicating very high agreement with the underlying data. The same model was additionally estimated using another dataset acquired on the same surface with the same camera, producing only negligible differences on the estimated polynomial parameters.

The correction of the measured distance $\hat{d}_k$ can be applied to each pixel k using

$$\hat{d}_{k,\text{corr}} = \hat{d}_k + d_{\text{dev}} \tag{3}$$

where $\hat{d}_{k,\text{corr}}$ is the corrected distance and d_{dev} is the correction output by the adapted measurement model and as computed shown in Section III-C. Once the correction is applied, a new set of Cartesian coordinates can be computed according to (1) to reconstruct the corrected 3-D point cloud.

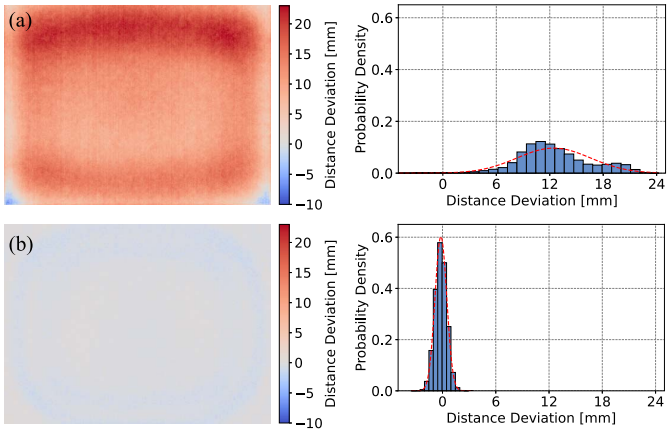


Fig. 11. Distance deviations before and after error compensation for a full image acquired at 55 cm using a second camera, i.e., a different one that was used for acquiring the dataset used for adapted model parameters. (a) Original depth image and its associated histogram. (b) Error-compensated depth image and its associated histogram. Both histograms also display normal distributions fit to the data with means of 12.5 and −0.2 mm and standard deviations of 4.1 and 0.8 mm.

This correction model was applied to a dataset acquired on the same surface with the second camera. An exemplary deviation image of the original dataset is shown in Fig. 11(a), where nonhomogeneously distributed deviations up to 25 mm are visible. The histogram of the original data shows that these deviations are neither normally distributed nor centered around zero. The dataset after applying the correction is shown in Fig. 11(b), where the deviations are bounded within ±2.5 mm and centered around zero as visible also on its corresponding histogram. Some weak systematics are still present in the corrected data, as visible on the residual circular pattern, likely showing small deviations between the systematic errors of both cameras and thus pointing to the limit of the error compensation capabilities; 83% of the deviations after correction, however, are actually below 1 mm and mainly defined by noise, as also supported by an approximately normal distribution visible on the histogram.

V. Conclusion

This work reports the procedures and results of an investigation on the distance measurement accuracy of a state-of-the-art ToF depth camera as driven by the most relevant sources of uncertainty that can be defined independently of the measured scene. The impact of self-warming, external temperature, noise, and sensor-related sources of systematic deviations with distance was analyzed experimentally, including interpixel and intercamera variability. The assessment was carried out using high-end metrology measurement setups to derive highly accurate and repeatable ground-truth data connected to the ICS of the sensor, thus providing reliable absolute uncertainty estimations.

The warm-up period, which takes about 40 min for the investigated camera, can introduce deviations on the measured distance of up to 2.5 mm. External temperature changes larger than 15 °C can induce deviations above 3 mm; however, if the external temperature remains stable within less than 2 °C, the deviations are only negligible compared to the range

noise floor. The correction of temperature-related errors using internal temperature readings was shown to be possible per individual cameras but impractical due to a lack of intercamera repeatability.

Range noise (2σ) is below 2 mm for distances up to 1 m and increases quadratically with distance reaching over 2 cm for ranges larger than 5 m. A noise increase with target reflectance was observed and as expected follows the expected behavior. Noise reduction by image averaging on a static target was shown to be efficient for acquisitions collected over several tens of seconds. Larger averaging was proven less effective likely due to higher noise spectral density on lower frequencies, and a precision limit at the level of a few hundred micrometers was identified.

Large, repeatable distance-related deviations reaching up to several centimeters were identified, being also systematic between different units of the same camera. Non-negligible interpixel deviations on those systematics up to several mllimeters were also observed, introducing nonhomogeneous average deviations across single depth images and suggesting the need for pixel-level error compensation.

The assessment served as a basis for giving operation and mitigation recommendations for noncorrectable deviations and exploring the compensation potential for systematic ones. The results showed room for large accuracy improvement of distance-related errors via measurement model adaptation on short ranges when the error budget is not dominated by noise. The systematics from one camera were used to derive a pixelwise adapted measurement model, which was then applied to measurements from another camera. The error compensation reduced the overall uncertainty about five times, with 80% of the corrected pixels yielding errors below 1 mm when the operation temperature remains within 2 °C. If larger temperature deviations are expected, intercamera error compensation still provides a significant improvement in relative accuracy and interpixel errors, although overlapped with temperature-dependent biases of ± 2.5 mm. The improvement potential demonstrated here is particularly relevant considering the intercamera transferability. Although proper error compensation measurements require extensive effort and high-end equipment as used in this work, the possibility of applying batch error compensation to several sensors based on a single measurement campaign may justify it in many applications. Those requiring accuracies on the level of 1 mm may particularly benefit from the outcomes of this investigation, including, e.g., small-scale deformation measurements of structural objects, 3-D shape reconstructions, or robotic manipulation tasks. A complete uncertainty budget evaluation, including uncertainty quantification and propagation for the error sources analyzed here as well as those related to specific application cases, will be addressed in the next steps of this research.

ACKNOWLEDGMENT

Robert Presl, Institute of Geodesy and Photogrammetry (IGP), ETH Zürich, Zürich, Switzerland, has built the mechanical setups for the experiments.

REFERENCES

[1] M. Faccio, E. Ferrari, M. Gamberi, and F. Pilati, "Human factor analyser for work measurement of manual manufacturing and assembly processes," *Int. J. Adv. Manuf. Technol.*, vol. 103, nos. 1–4, pp. 861–877, Jul. 2019.

[2] M. Ochman, M. Skoczeń, D. Krata, M. Panek, K. Spyra, and A. Pawłowski, "RGB-D odometry for autonomous lawn mowing," in *Proc. Int. Conf. Artif. Intell. Soft Comput.* Cham, Switzerland: Springer, 2021, pp. 81–90.

[3] M. Yan, Z. Li, X. Yu, and C. Jin, "An end-to-end deep learning network for 3D object detection from RGB-D data based on Hough voting," *IEEE Access*, vol. 8, pp. 138810–138822, 2020.

[4] A. Haleem and M. Javaid, "3D scanning applications in medical field: A literature-based review," *Clin. Epidemiol. Global Health*, vol. 7, no. 2, pp. 199–210, 2018.

[5] Z. Zhang, Z. Cui, C. Xu, Z. Jie, X. Li, and J. Yang, "Joint task-recursive learning for RGB-D scene understanding," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 10, pp. 2608–2623, Oct. 2020.

[6] B. Langmann, K. Hartmann, and O. Loffeld, "Depth camera technology comparison and performance evaluation," in *Proc. ICPRAM*, vol. 2, 2012, pp. 438–444.

[7] A. C. Ulku *et al.*, "A 512×512 SPAD image sensor with integrated gating for widefield FLIM," *IEEE J. Sel. Topics Quantum Electron.*, vol. 25, no. 1, Jan./Feb. 2018, Art. no. 6801212.

[8] *Time-of-Flight Forged Ahead: Design Tips to Boost 3D Performance and Cut Integration Time and Cost (V.1.3)*, Lucid Vision Labs, Richmond, BC, Canada, 2019.

[9] V. Frangez, E. Lloret-Fritschi, N. Taha, F. Gramazio, M. Kohler, and A. Wieser, "Depth-camera-based rebar detection and digital reconstruction for robotic concrete spraying," *Construct. Robot.*, vol. 5, pp. 191–202, Oct. 2021.

[10] V. Frangez, D. Salido-Monzú, and A. Wieser, "Surface finish classification using depth camera data," *Autom. Construction*, vol. 129, Sep. 2021, Art. no. 103799.

[11] M. Carfagni *et al.*, "Metrological and critical characterization of the Intel D415 stereo depth camera," *Sensors*, vol. 19, no. 3, p. 489, Jan. 2019.

[12] *Helios HLS003S Technical Reference (1.18.0.0)*, Lucid Vision Labs, Richmond, BC, Canada, 2020.

[13] R. Horaud, M. Hansard, G. Evangelidis, and C. Ménier, "An overview of depth cameras and range scanners based on time-of-flight technologies," *Mach. Vis. Appl.*, vol. 27, no. 7, pp. 1005–1020, Oct. 2016.

[14] M. Zámečníková, A. Wieser, H. Woschitz, and C. Ressl, "Influence of surface reflectivity on reflectorless electronic distance measurement and terrestrial laser scanning," *J. Appl. Geodesy*, vol. 8, no. 4, pp. 311–326, Jan. 2014.

[15] E. Lachat, H. Macher, T. Landes, and P. Grussenmeyer, "Assessment and calibration of a RGB-D camera (Kinect v2 sensor) towards a potential use for close-range 3D modeling," *Remote Sens.*, vol. 7, no. 10, pp. 13070–13097, 2015.

[16] M. Lindner, I. Schiller, A. Kolb, and R. Koch, "Time-of-flight sensor calibration for accurate range sensing," *Comput. Vis. Image Understand.*, vol. 114, no. 12, pp. 1318–1328, 2010.

[17] T. Kahlmann, "Range imaging metrology: Investigation, calibration and development," Ph.D. dissertation, Dept. Civil, Environ. Geomatic Eng., ETH Zürich, Zürich, Switzerland, 2007.

[18] J. Jung, J. Y. Lee, Y. Jeong, and I. S. Kweon, "Time-of-flight sensor calibration for a color and depth camera pair," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 7, pp. 1501–1513, Jul. 2015.

[19] P. Yu, S. Yang, and S. Chen, "Accuracy improvement of time-of-flight depth measurement by combination of a high-resolution color camera," *Appl. Opt.*, vol. 59, no. 35, pp. 11104–11111, 2020.

[20] V. Frangez, D. Salido-Monzú, and A. Wieser, "Depth-camera-based in-line evaluation of surface geometry and material classification for robotic spraying," in *Proc. 37th Int. Symp. Autom. Robot. Construct. (ISARC)*, Oct. 2020, pp. 693–702.

[21] S. Pasinetti, M. Hassan, J. Eberhardt, M. Lancini, F. Docchio, and G. Sansoni, "Performance analysis of the PMD camboard picoflexx time-of-flight camera for markerless motion capture applications," *IEEE Trans. Instrum. Meas.*, vol. 68, no. 11, pp. 4456–4471, Jan. 2019.

[22] A. A. A. Beshr and I. M. Abo Elnaga, "Investigating the accuracy of digital levels and reflectorless total stations for purposes of geodetic engineering," *Alexandria Eng. J.*, vol. 50, no. 4, pp. 399–405, Dec. 2011.

[23] D. D. Lichti and J. Lampard, "Reflectorless total station self-calibration," *Surv. Rev.*, vol. 40, no. 309, pp. 244–259, Jul. 2008.

[24] C. Holst, H. Neuner, A. Wieser, T. Wunderlich, and H. Kuhlmann, "Calibration of terrestrial laser scanners," *Allgemeine Vermessungs-Nachrichten, AVN, Zeitschrift für alle Bereiche der Geodäsie und Geoinf.*, vol. 123, no. 6, pp. 147–157, 2016. [Online]. Available: https://hdl.handle.net/20.500.11811/8764

[25] A. P. P. Jongenelen, D. G. Bailey, A. D. Payne, A. A. Dorrington, and D. A. Carnegie, "Analysis of errors in ToF range imaging with dual-frequency modulation," *IEEE Trans. Instrum. Meas.*, vol. 60, no. 5, pp. 1861–1868, May 2011.

[26] S. H. Lee, S. C. Kwon, H. B. Chae, J. Y. Park, H. J. Kang, and J. D. K. Kim, "Digital hologram generation for a real 3D object using by a depth camera," in *Proc. J. Phys., Conf.*, 2013, vol. 415, no. 1, Art. no. 012049.

[27] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 11, pp. 1330–1334, Nov. 2000.

[28] P. Artaso and G. López-Nicolás, "Volume estimation of merchandise using multiple range cameras," *Measurement*, vol. 89, pp. 223–238, Jul. 2016.

[29] M. Georgiev, R. Bregović, and A. Gotchev, "Fixed-pattern noise modeling and removal in time-of-flight sensing," *IEEE Trans. Instrum. Meas.*, vol. 65, no. 4, pp. 808–820, Apr. 2016.

[30] P. E. Ciddor and R. J. Hill, "Refractive index of air. 2. Group index," *Appl. Opt.*, vol. 38, no. 9, pp. 1663–1667, 1999.

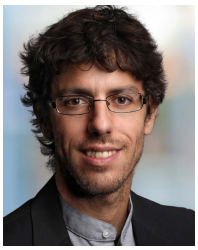
[31] M. Frukacz, R. Presl, A. Wieser, and D. Favot, "Pushing the sensitivity limits of RTS-based continuous deformation monitoring of an Alpine valley," *Appl. Geomatics*, vol. 9, no. 2, pp. 81–92, Jun. 2017.

[32] H. Vennegeerts, E. Richter, J. Paffenholz, H. Kutterer, and M. Hennes, "Genauigkeitsuntersuchungen zum kinematischen einsatz terrestrischer laserscanner," in *Proc. AVN*, 2010, pp. 140–147.

[33] R. Y. Tsai and R. K. Lenz, "A new technique for fully autonomous and efficient 3D robotics hand/eye calibration," *IEEE Trans. Robot. Autom.*, vol. 5, no. 3, pp. 345–358, Jun. 1989.

[34] Y. Sakamoto, M. Ishiguro, and G. Kitagawa, *Akaike Information Criterion Statistics*, vol. 81, no. 10, D. Reidel, Ed. Dordrecht, The Netherlands: Taylor & Francis, 1986, p. 26853.

Valens Frangez is currently a Scientific Assistant with the Group of Geosensors and Engineering Geodesy, ETH Zürich, Zürich, Switzerland. In August 2019, he joined the interdisciplinary field of NCCR Digital Fabrication as a Ph.D. Researcher. His main research interests are industrial metrology and machine vision for an in-line 3-D shape determination and geometric assessment of robotically fabricated elements.



David Salido-Monzú is currently a Post-Doctoral Researcher with the Chair of Geosensors and Engineering Geodesy, ETH Zürich, Zürich, Switzerland, where he coordinates research on laser-based dimensional metrology and material probing. His main interests revolve around optical metrology, sensing, and electronic instrumentation, currently focusing on the application of optical frequency combs to enhance distance measurement technology and remote spectroscopy.

Andreas Wieser has been a Full Professor of geosensors and engineering geodesy with ETH Zürich, Zürich, Switzerland, since 2012. He has held various academic positions at the Vienna University of Technology, Vienna, Australia; the Graz University of Technology, Graz, Austria; and the University of Calgary, Calgary, AB, Canada. He worked as a Product Manager for GPS-based tolling in industry before. His research interests are geodetic monitoring, digitization of reality, ubiquitous positioning, parameter estimation, and quality control.