Daniel Janos[1], Łukasz Ortyl[2], Przemysław Kuras[3]

Low-Cost 3D Depth Sensors for Mobile Applications and Control Systems – Accuracy Assessments Using Surveying Techniques

Abstract: This article focuses on low-cost LiDAR (light detection and ranging) sensors and 3D depth cameras. Particular attention was paid to their accuracy and compliance with the technical specifications that were provided by their respective manufacturers. The following devices were tested: Stereolabs ZED 2i, Stereolabs ZED, and Intel RealSense D435i depth cameras, and the Intel RealSense L515 LiDAR sensor. An experiment was carried out to measure a geometrically diverse environment (which is typical for in-motion imaging) where both the measurement range and the distortion that is generated by each device's algorithms on edges, folds, planes, and 3D objects could be evaluated. Depth sensors are often used with excessive confidence as to their geometric reliability. The aim of this work is to assess the actual accuracy of such sensors, which may constitute the ground truth for accuracy losses that could result from the operations of autonomous vehicles. Based on the results, the accuracy information that was provided by the respective manufacturers was difficult to obtain under real conditions. It was found that the low-cost devices could be used in industrial projects, but their operations must take place under certain conditions and settings. It was also necessary to know their capabilities and limitations in order to take full advantage of what they offer.

Keywords: low-cost sensors, depth camera, LiDAR, comparison, point cloud accuracy

[1] AGH University of Krakow, Faculty of Geo-Data Science, Geodesy, and Environmental Engineering, Krakow, Poland, email: janos@agh.edu.pl (corresponding author), https://orcid.org/0000-0002-5394-7922

[2] AGH University of Krakow, Faculty of Geo-Data Science, Geodesy, and Environmental Engineering, Krakow, Poland, email: ortyl@agh.edu.pl, https://orcid.org/0000-0002-9151-7931

[3] AGH University of Krakow, Faculty of Geo-Data Science, Geodesy, and Environmental Engineering, Krakow, Poland, email: kuras@agh.edu.pl, https://orcid.org/0000-0001-5066-6220

1. Introduction

1.1. Solid-State LiDARs

Over the past few years, low-cost LiDAR sensors have become extremely popular for many applications that are related to the movement automation of mobile platforms. They have not achieved the performance that is inherent in high-end mechanical laser scanners, but they have met sufficient requirements in many applications. As Raj et al. [1] noted, "despite its weaknesses, the low cost of this 3D LiDAR system makes it an attractive solution for applications such as robotics, surveying, agriculture, and education." Of the many existing scanning mechanisms, they saw great potential in low-cost solutions.

Solid-state LiDAR sensors are used in SLAM (simultaneous localization and mapping) solutions. Nam and Gon-Woo [2] presented the fact that, despite their meager costs, they could achieve high performance and achieve the results that were inherent in expensive mechanical devices. In their research, they used Livox Horizon and Intel RealSense L515 sensors. Commercially available SLAM devices such as ZEB-REVO also use a 2D laser scanning device (Hokuyo UTM-30LX-F) [3]. Similar research conclusions were drawn by Wei et al. [4]; the authors pointed to the increasing role of solid-state LiDAR on the LiDAR market. As part of their work, they constructed a 2D-3D mapping system using Livox Mid-40 and Hokuyo UST-20LX LiDAR sensors with varying FOVs (fields of view). The device was successfully tested for feature-detection in urbanized spaces.

LiDARs are fundamental sensors in many robotic applications – particularly in autonomous driving [5]. Kutila et al. [6] noted that research was currently underway to select optimal devices and their parameters for self-driving cars – especially under adverse weather conditions. In contrast, Li et al. [7] presented the potential of using a solid-state LiDAR sensor; its price was significantly lower than that of conventional sensors. The authors applied a Livox Horizon device to develop a real-time method of mapping and inertial odometry. The 3D LiDAR streams were synchronized with a six-axis IMU (inertial measurement unit) in order to obtain a globally consistent map. Another LiDAR sensor, Ouster OS1-16, was applied as a component of an autonomous robot that worked in vineyards [8]. A sensor with an appropriate operating range allowed it to provide data that was necessary for recognizing vine rows. Moreover, Haddeler et al. [9] proved that data that was obtained with LiDAR sensors could create the basis for combining data from many autonomous vehicles during their operations in multi-robot mode. Another application is low-cost heritage documentation; Murtiyoso et al. [10] evaluated an Apple iPad Pro solid-state LiDAR sensor by comparing its results with a reference point cloud. They found outliers of the order of 35 or 7.5% depending on the lighting, the textural properties, and the level of object complexity and formulated several useful recommendations for this sensor type when used for the architectural inventories of historic buildings.

The topic of assessing the accuracy of laser scanners has been known since the devices became widely available [11]. Similar research has been conducted on low-cost scanners. One of the first efforts was the test that was carried out by Lee and Ehsani [12] for the Hokuyo URG-04LX 2D sensor by comparing it with the Sick LMS200 scanner. Since then, the topic of research on low-cost LiDAR accuracy has appeared frequently in the literature. These devices can be tested from a number of aspects; e.g., drift, surface color, surface material, incident angle, luminosity, and distance. Tests have been conducted for both commercially available sensors (e.g., Slamtec RPLidar A1-M8, Hokuyo URG-04LX, and Hokuyo UTM-30LX-EW [13]) and prototype sensors (e.g., a 905-nm-wavelength laser-based LiDAR sensor for autonomous vehicles [14]). The most common way to assess the accuracy of LiDAR sensors is through comparative tests with a reference point cloud that is obtained from a TLS scan. An example of such a study for the Velodyne HDL-32E sensor was presented by Jozkow et al. [15].

The Intel RealSense L515 LiDAR sensor was used as part of the research that was conducted in our thesis. For this device, Lam et al. [16] reported RMSE values within a range of 2–6 mm, which was determined to be relative to a metrology-grade reference value. However, it should be taken into account that these were determined by observing an object that was covered with white non-reflective paint from a distance of 0.5 m. The noise that was determined for this device was evaluated within a range of 1–3 mm. In contrast, Plaß et al. [17] presented the results of an accuracy assessment for this sensor based on a comparison with a professional geodetic scanner for a larger area – in a 40 m^2 room. In this case, a comparison with the geodetic measurement gave an average difference of 8 mm (based on a comparison of the distances to the reference targets) and 13 mm (by a comparison of the reference lengths); the measurement noise was set at 4 mm. However, other values were obtained by Breitbarth et al. [18] when checking the accuracy of the L515 sensor – particularly in terms of the actual working range and the working performance that were offered for each resolution. They summarized their study with standard deviation values of 13 mm for the distance that was determined between the spheres and 22 mm for flatness-measurement deviation. They concluded that the results that were acquired by the L515 sensor were more reliable when compared to the widely used depth cameras (on the example of the Intel D415 camera, which delivered more unstable results of the measurement).

1.2. Depth Cameras

Depth cameras provide an alternative solution. According to Zollhöfer [19], "an RGB-D sensor is the combination of a conventional color camera (RGB) with such a depth sensor (D)." Their most common applications are human pose/gesture recognition, object recognition (e.g., species and plant parts), 3D modeling, autonomous flight control, and robot navigation. For RGB-D cameras, inaccuracy is considered in two categories: first, from the pose estimation of a camera; and second, from the inaccuracy of the camera itself [20].

The latter contains the following:
– calibration errors,
– random noise,
– camera sensor limitations (resolution),
– motion blur effects.

As the sensors and measured objects were stationary in the presented study, this paper focuses on assessing the precision and accuracy of those RGB-D sensors that were not affected by motion blur.

The first of the depth cameras that was investigated in our work was the LiDAR RealSense D435i by Intel. Accuracy tests for this type of device were carried out by Carfagni et al. [21] using another model (D415) as an example. They found a high agreement of the imaged plane at 2.5 mm relative to the ground truth, with this value determined for a range from 150 to 500 mm. On the other hand, Moghari et al. [22] indicated the possibility of achieving accuracy levels of 1–2% of the distance from a sensor to an object; however, the observations were only conducted from a distance of 0.35 m and concerned observations of human head motions.

A comparative test of several sensors was carried out by Curto and Araujo [23]. Of the devices that were tested (Intel RealSense L515, SR305, and D415), they chose the last one as having the best behavior in terms of its statistical stability (expressed as repeatability or precision). It should be noted that the tests were mainly related to assessing the repeatability (precision) according to the type of surface that was observed. In contrast, Lourenço and Araujo [24] chose the same set of sensors for testing; in terms of accuracy and precision, the L515 sensor achieved the best results. In addition to this device having the lowest average error, its correct operation was also least-affected by distance; its standard deviation as a measure of precision was less than 0.5 mm for distances of up to 3.5 m.

A similar set of devices was tested by Servi et al. [25] with one exception: the SR305 was replaced by the D455. Their results varied for the individual sensors. For a calibrated sphere at a very close range, the best results were obtained by the D415. In contrast, probing errors and distortion characteristics (determined according to ISO 10360-13) did not indicate an advantage for any of the sensors, nor did the test that involved 3D object reconstruction. However, the L515 had a clear advantage over the others in determining the systematic depth error for a plane at distances of 0.5–1.5 m.

The other devices that were tested in our study were the Stereolabs ZED and ZED 2i cameras. In the research on these devices, however, the concept of accuracy was used in the context of image recognition rather than expressing the geometric quality of the acquired point clouds. For example, Tadic [26] applied the ZED 2i sensor for charging-socket detection in electric vehicles. Conversely, Connolly et al. [27] used the cameras to improve the detection rate during visual inspections that were conducted from unmanned aerial vehicles.

Research on the depth-error modeling of the ZED sensor (the predecessor to the ZED 2i) was conducted by Ortiz et al. [28]. A mathematical model of the depth error that was determined when using this sensor was proposed; it showed that the RMS in-depth error was related to the camera resolution. In addition, the authors found that at the highest resolution (a range of 20 m) was achievable, although an acceptable error rate was met for shorter distances in some applications.

1.3. Motivation

The presented literature research shows that low-cost LiDAR sensors and depth cameras are very popular for positioning autonomous robots and for their avoidance of obstacles [29, 30]. Combining these two types of sensors particularly requires the development of efficient self-calibration algorithms [31, 32]. The sensors are widely used in scientific projects for spatial-data detection, creating 3D models of objects, the route-planning of autonomous systems, human and object recognition, etc. [33–36]. However, the data that was provided by these devices could not be treated uncritically – especially in terms of accuracy and fidelity in the imaging reality. Furthermore, it should be noted that the sensors were used under real conditions in many studies (especially in motion), while less attention was paid to the loss of accuracy as compared to the reference values.

Such a critical approach to sensor accuracy was presented by Holder et al. [37]. To assess the reliability of the reference accuracy of vehicle positioning using GNSS sensors, the authors used the concept of "super-reference"; this means the best-available measurement to achieve the ground truth. The research that was carried out in their article was used to assess the actual accuracies and ranges of the low-cost depth sensors; they checked whether the technical specifications and manufacturers' assurances were confirmed in reality. The sensors that were selected for their study had not yet been compared in terms of their geometric accuracies when the reference value was provided by a professional geodetic laser scanner. The aim of our study was to recognize the capabilities and limitations of LiDAR sensors and depth cameras, together with recognitions of their practical applicabilities.

1.4. Article Structure

The remainder of this paper is arranged as follows. The methodology of the conducted research is presented in Section 2, while a presentation of the equipment that was used for the study is included in Section 3, along with the measurement setup on the test field. An explanation of the procedure that was used for the data processing is presented in Section 4, along with the sample results. Detailed results of the sensor precision and accuracy tests are discussed in Section 5; since numerical values do not express all of the observed problems, a visual assessment of the acquired point clouds is also illustrated there. The discussion is presented in Section 6, and the final conclusions are given in Section 7.

2. Methodology

When developing the methodology to address the research problem, the aim was to provide answers to the following four questions:

1) What was done to answer the research question?
   – four popular and affordable sensors were selected for mapping environment and generating point clouds;
   – quality and accuracy of acquired clouds were checked;
   – results that were obtained by selected devices were compared with each other.
2) How it was done?
   – inspections were carried out using precise geodetic instruments (two laser scanners, and one robotic total station), which were at least one-order-of-magnitude more accurate than tested sensors;
   – static measurements were taken with all devices in same environment from same measuring station under uniform conditions;
   – obtained results were compared by calculating cloud-to-cloud (C2C) distances in different configurations;
   – visual assessments of point clouds that were generated by sensors were carried out.
3) How can the experiment be substantiated?
   – design of experiment allowed for all devices to be tested under uniform conditions;
   – prepared measurement environment ensured actual measurement ranges of sensors and their abilities to map various types of objects, and accuracy of generated point clouds could be verified.
4) How were the results analyzed?
   – numerical calculations of C2C distances were performed to present noise of point clouds of tested sensors as well as their deviations from reference cloud;
   – visual evaluations of point clouds allowed for mapping quality of the measured objects on edges and planes to be presented in tangible manner.

3. Equipment and Test Field

This section briefly describes the equipment (Fig. 1) that was used to perform the experiment; the measurement environment has been addressed, as have the accessories that were necessary for reliable testing. The needed software, measurement, and data-processing techniques that enabled the validation task to be carried out correctly have also been presented.

Fig. 1. Sensors used in tests as well as necessary equipment and accessories
(from left to right): Leica MS50 robotic total station; Z+F 5010C laser scanner;
Leica C10 laser scanner; Stereolabs ZED 2i, Stereolabs ZED, and Intel D435i depth cameras;
Intel L515 LiDAR camera

3.1. ZED 2i Depth Camera by Stereolabs

The ZED 2i is a depth camera that was developed by Stereolabs that utilizes stereo vision to capture 3D images of the surrounding environment [27, 38]. This technology works by using two cameras to take images of the same scene from slightly different angles; these are then used to calculate the depths and distances of the objects in the scene. This allows the ZED 2i to measure the sizes and locations of objects in its field of view in real time, perceive their surroundings in three dimensions, and navigate around the objects. This is also useful for autonomous vehicles – it can help a vehicle understand its surroundings and make decisions about where to go and how to avoid obstacles. In addition, the ZED 2i can be used in augmented-reality applications – it can deliver a realistic spatial representation of an environment to be overlaid with virtual objects.

A computer/CPU (central processing unit) with a minimum 4 GB RAM (random-access memory), dual-core 2.3 GHz processor, and NVIDIA GPU (graphics processing unit) with compute capability >3.0 is required to support the camera's SDK (software development kit). The lack of such a setup (or a more powerful configuration) may prevent the device from working in real time in projects with limited spaces, weights, or powers (e.g., UAVs – unmanned aerial vehicles), and the collected data will require post-processing. This significantly increases the cost of the whole set-up, as a powerful computer is required in addition to the camera. However, the idea of data processing on an external computer seems to be a valid one for several reasons. First, the camera itself does not heat up and can, therefore, be enclosed in a sealed housing. What is more, the parameters of the computer can be adjusted to suit the needs of the project. As the manufacturer is constantly improving the SDK (especially the AI [artificial intelligence] mode, which could be made impossible in the future by an outdated CPU if it were pre-installed in the camera), the camera itself "ages" much more slowly.

Thanks to the aforementioned SDK, the user can access generated data such as a point clouds, depth maps, RGB images, or ambient sensor data. This data can be accessed through the manufacturer's dedicated programs within the SDK or through the support of a variety of programming languages and platforms, such as C++, Python, and ROS (Robot Operating System) in the user's own applications.

The ZED 2i camera is the latest item from the devices that were tested in this article; it represents the best value-for-money proposition on the market today in its price range (around $500 [excluding CPU]). It is therefore worth checking whether newer algorithms, manufacturing processes, and technology go hand-in-hand with increases in point-cloud-measurement accuracy. This information will be particularly useful for projects that involve the use of this camera in order to evaluate its usefulness.

Some of the key technical parameters of the Stereolabs ZED 2i camera can be seen in Table 1. Of all of the devices that were tested, it by far offers the most to the user in terms of field of view, range, refresh rate, additional information from ambient sensors, and robustness under environmental conditions. Overall, the Stereolabs ZED 2i is a depth camera that is relevant for a broad range of applications; it has the ability to capture high-resolution 3D images in real-time, a compact size, and a lightweight design.

Table 1. Main parameters of Stereolabs ZED 2i, Stereolabs ZED, Intel D435i, and Intel L515 sensors

Parameter	Sensor			
	Stereolabs ZED 2i	Stereolabs ZED	Intel D435i	Intel L515
Type	RGB-D camera	RGB-D camera	RGB-D camera	LiDAR
Depth field of view	horizontal 110°, vertical 70° (standard 2.1 mm lenses)	horizontal 90°, vertical 60°	horizontal 87°, vertical 58°	horizontal 70°, vertical 55°
Maximum resolution	2208 × 1242 px @ 15 FPS (frames per second), up to 100 FPS on lower resolutions	2208 × 1242 px @ 15 FPS, up to 100 FPS on lower resolutions	1280 × 720 px @ 30 FPS, up to 90 FPS on lower resolutions	1024 × 768 px @ 30 FPS
Depth range	0.3–20 m (standard 2.1 mm lenses)	0.5–25 m	0.3–3 m	0.25–9 m
Depth accuracy (declared by manufacturers)	<1% up to 3 m, <5% up to 15 m	<2% up to 3 m, <4% up to 15 m	<2% @ 2 m	±5 mm @ 1 m, ±14 mm @ 9 m

Source: [39–45]

3.2. ZED Depth Camera by Stereolabs

The Stereolabs ZED camera is the first generation of depth cameras from this manufacturer [28, 46]; it was included in this compilation in order to test the real differences and progress in the point-cloud generation between it and the manufacturer's latest offering.

Some of the key technical parameters of the Stereolabs ZED depth camera are presented in Table 1; it had a slightly worse specification than ZED 2i. Several features were missing here, such as weather resistance, a detachable cable (which was the cause of a transmission interruption in the test unit when the cable was in an unfortunate position), or IMU, temperature and pressure sensors (which the manufacturer improved in the newer model). Due to the slightly different focal lengths of the lenses, its field of view, declared range and measurement accuracy were different from the ZED 2i.

3.3. RealSense D435i Depth Camera by Intel

Developed by Intel, the RealSense D435i is a stereoscopic RGB-D camera; it is equipped with a RealSense Vision D4 processor. In addition to a stereo vision module that is composed of two depth sensors, it is equipped with an infrared projector and an RGB camera [47, 48]. An additional element of the D435i sensor is an IMU module that can be used to detect rotation and motion in six degrees of freedom. This model is particularly popular for smaller projects due to its affordable price (around $350), IMU equipment, small size, no requirement for data-processing on an external PC (personal computer), and good documentation and support from the manufacturer's SDK. The last feature allows users to easily create their own applications in the C/C++, C#/.NET, Matlab, Node.js, or Python programming languages using Intel's depth cameras and LiDAR sensors.

Some of the key technical parameters of the Intel RealSense D435i depth camera can be seen in Table 1. RealSense is definitely a much smaller camera than its Stereolabs competitors; however, it finds uses in many projects where range is not crucial but size, weight, price, and/or the lack of a powerful central processing unit are. In later tests, the claimed accuracy and measurement range were compared with the technical specifications.

3.4. RealSense L515 LiDAR by Intel

Also developed by Intel, the RealSense L515 is a solid-state LiDAR (light detection and ranging) sensor that is mainly designed for use in robotics applications [2, 49]. Some of its key features are its small size and low power consumption. The sensor is just 61 mm in diameter and 26 mm in width, making it well-suited for use in small vehicles and robots. The L515 also has a fairly wide field of view and high resolution, which are necessary for creating detailed 3D maps.

Another feature of the Intel L515 LiDAR is its ability to work under a variety of lighting conditions – even in complete darkness. It also has a built-in shutter that can be used to block out sunlight, which helps improve its performance under bright conditions. Unfortunately, the correct operation of the device may be impaired under daylight conditions due to the infra-red light that is emitted by the Sun; therefore, using the camera outdoors may reduce the quality of the depth images. Due to its power efficiency, the L515 LiDAR sensor performs best under controlled lighting conditions and indoors [50]. In terms of performance, the L515 LiDAR features high accuracy (±5 mm @ 1 m, ±14 mm @ 9 m) and an update rate of 30 Hz [39]. This update rate is fast enough for most robotic applications.

3.5. Acquisition of Reference Data

The reference data set (i.e., a single point cloud that was acquired with a Leica C10 professional geodetic scanner) was used as the ground truth for the results of the examined depth cameras and LiDAR sensor. The Leica C10 had a relatively low vertical distance from the rotating mirror axis to the base (tribrach) of the scanner; this was extremely advantageous when designing and printing the special mounts for each camera (Fig. 1), enabling measurements with the scanner tribrach at exactly the same height and on the same plumb line.

The accuracy characteristics of the scanner are shown in Table 2. The accuracy of the 3D position was declared by the manufacturer as being ±6 mm @ 50 m. In contrast, the Intel L515 (which was theoretically the most accurate of the sensors that were tested) claimed ±14 mm @ 9 m in its specifications. Although the Leica C10 scanner has already been established as a proven device by many experts [51, 52] and achieved higher-than-declared accuracies at our tested (short) distances, the results were further checked with another geodetic scanner – the Z+F 5010C (data in Table 2).

Table 2. Selected measurement-accuracy parameters of equipment used to acquire reference point cloud and verify it – Leica C10 and Z+F 5010C scanners, and Leica MS50 total station

Parameter	Leica C10	Z+F 5010C	Leica MS50	Unit	Source
Distance accuracy @ 50 m	4	2.2	2.1	mm	datasheet [53–55]
Angle accuracy [Hz/V]	12/12	25/25	1/1	"	datasheet [53–55]
Position accuracy @ 50 m	4.9	6.4	2.1	mm	calculated

In order to link the measurements as accurately as possible, all of the scanning targets were measured with a Leica MS50 total station using the reflector-less technique. The point-cloud registration (linking) was performed in Leica Cyclone software; the maximum residual for registering the Leica C10 to the Leica MS50

was 3 mm (with an average of 1 mm). In turn, the residual of the Z+F 5010C to the Leica MS50 was 2 mm (with an average of 1 mm). For the combined registration, the maximum error of the Leica C10 to the Z+F 5010C was 5 mm (with an average of 1 mm); this allowed for the results that were obtained with the Leica C10 to be considered to be appropriate for reference purposes.

3.6. Experimental Setup

The measurements were taken in a lecture room that offered sufficient space and suitable facilities for checking the range and accuracy of the sensors that were mentioned earlier. Together with professional accessories for laser scanning and photogrammetry (standard targets for scan registrations), the walls, floor, corners, and furnishings of the room were used to correctly orient the point clouds in relation to the reference scans and check the overall quality of the measurements. Sensitive areas (e.g., the edges of planes, changes in color, object textures, spatial depths, etc.) received particular attention, as were their alignments to the reference scan that was obtained with the Leica C10. The measurement stand and its configuration are shown in Figure 2.

Fig. 2. Measurement targets that were visible from measuring station

To facilitate the correct orientation of the test scans in relation to the reference cloud, the sensors were mounted coaxially with the Leica C10 scanner and at the same height thanks to 3D-printed adapters (Fig. 1). The adapter print accuracy oscillated within ±0.2 mm; they were designed in such a way that the origins of the XYZ sensor coordinate systems remained in the same place regardless of the rotation of the sensors on the carrier pin that was mounted on the geodetic tribrach.

3.7. Performing Experiment

The experiment started with reference measurements by the Leica MS50 robotic total station.

Reflector-less measurements were taken for all of the flat scanning targets; the obtained coordinates were used to fit two point clouds from professional geodetic scanners: the Z+F 5010C, and the Leica C10. The main ground-truth data to which the measurements of the depth cameras under examination would be related was that from the Leica C10; the measurement with the Z+F scanner was solely a control for the Leica C10 measurement.

Full 360° scans were taken together with photographs in order to obtain a colored point cloud. The scanning resolution was set to 10 mm @ 10 m (horizontally and vertically), which allowed us to obtain a dense point cloud at such a short distance (indoor measurements). It was possible to capture both the previously mentioned flat walls and all of the other measurement targets from a single position.

The depth cameras were tested consecutively – each separately toward the field with targets.

The first was the Stereolabs ZED 2i; measurements were taken with it at all of its possible resolutions (2K, HD1080, HD720, and VGA). Each lasted a few seconds, and the results were raw files in the.svo format; these were successively used to prepare point clouds in the software that was provided by the manufacturer in the SDK package (ZED Depth Viewer). In the cases of the Stereolabs cameras, the raw files were comprised of stereo images that were converted into point clouds in the software. This allowed for some of the point-cloud-generation parameters to be adjusted later on. The post-processing will be described in more detail in the next section.

After changing the camera on the carrier pin, a measurement was performed with the Stereolabs ZED camera. Similar to the Stereolabs ZED 2i, the raw.svo files from all of its possible resolutions (2K, HD1080, HD720, and VGA) were saved for the post-processing.

Measurements were then taken with the Intel D435i camera; this device offered six resolutions (1280 × 720, 848 × 480, 640 × 480, 640 × 360, 480 × 270, and 424 × 240). At each of these resolutions, a few seconds of measurements were stored in.bag files. In this case, however, these files already contained generated point clouds – a recording/video of the measurements. Moreover, changing the parameters of the point-cloud calculations in post-processing was impossible; it was only possible to select any frame of the video from the.bag file and save it (e.g., in the.ply format). For this reason, the field measurements were taken by using the different modes that were offered by the camera: default, hand, high accuracy, high density, and medium density.

Similarly, measurements were taken with the Intel L515 LiDAR, which also provided final point cloud frames (with resolutions of 1024 × 768, 640 × 480, and 320 × 240);

these were saved directly to the.bag file. For this sensor, the available modes were no ambient light, low ambient light, max range, and short range. Comparisons of all of the modes will be presented in the following sections.

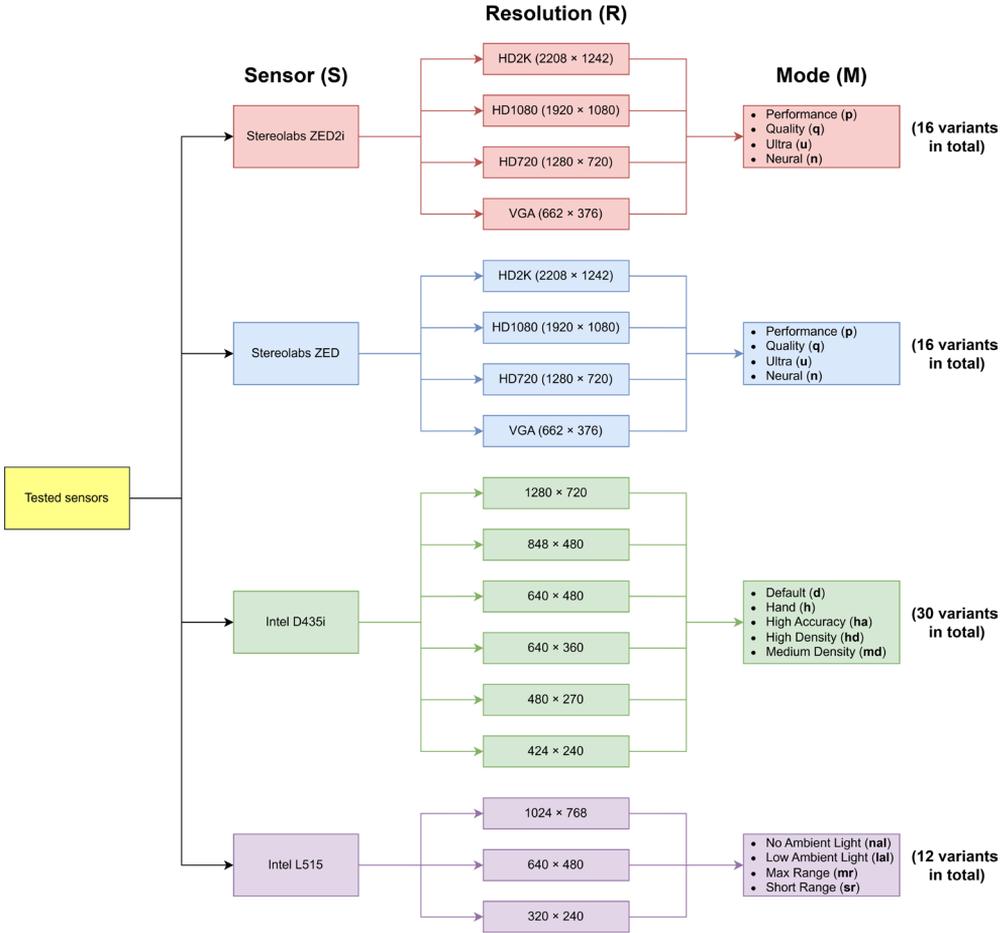All of the resolution and mode combinations for each camera are shown in Figure 3.

Resolution (R)

Sensor (S)

Mode (M)

Stereolabs ZED2i
- HD2K (2208 × 1242)
- HD1080 (1920 × 1080)
- HD720 (1280 × 720)
- VGA (662 × 376)
- Performance (**p**)
- Quality (**q**)
- Ultra (**u**)
- Neural (**n**)

(16 variants in total)

Stereolabs ZED
- HD2K (2208 × 1242)
- HD1080 (1920 × 1080)
- HD720 (1280 × 720)
- VGA (662 × 376)
- Performance (**p**)
- Quality (**q**)
- Ultra (**u**)
- Neural (**n**)

(16 variants in total)

Tested sensors

Intel D435i
- 1280 × 720
- 848 × 480
- 640 × 480
- 640 × 360
- 480 × 270
- 424 × 240
- Default (**d**)
- Hand (**h**)
- High Accuracy (**ha**)
- High Density (**hd**)
- Medium Density (**md**)

(30 variants in total)

Intel L515
- 1024 × 768
- 640 × 480
- 320 × 240
- No Ambient Light (**nal**)
- Low Ambient Light (**lal**)
- Max Range (**mr**)
- Short Range (**sr**)

(12 variants in total)

Fig. 3. Tree diagram of resolutions and modes
in which measurement data was collected

4. Data Processing

The data that was obtained in accordance with the assumed methodology was processed as is shown in the pipeline in Figure 4.
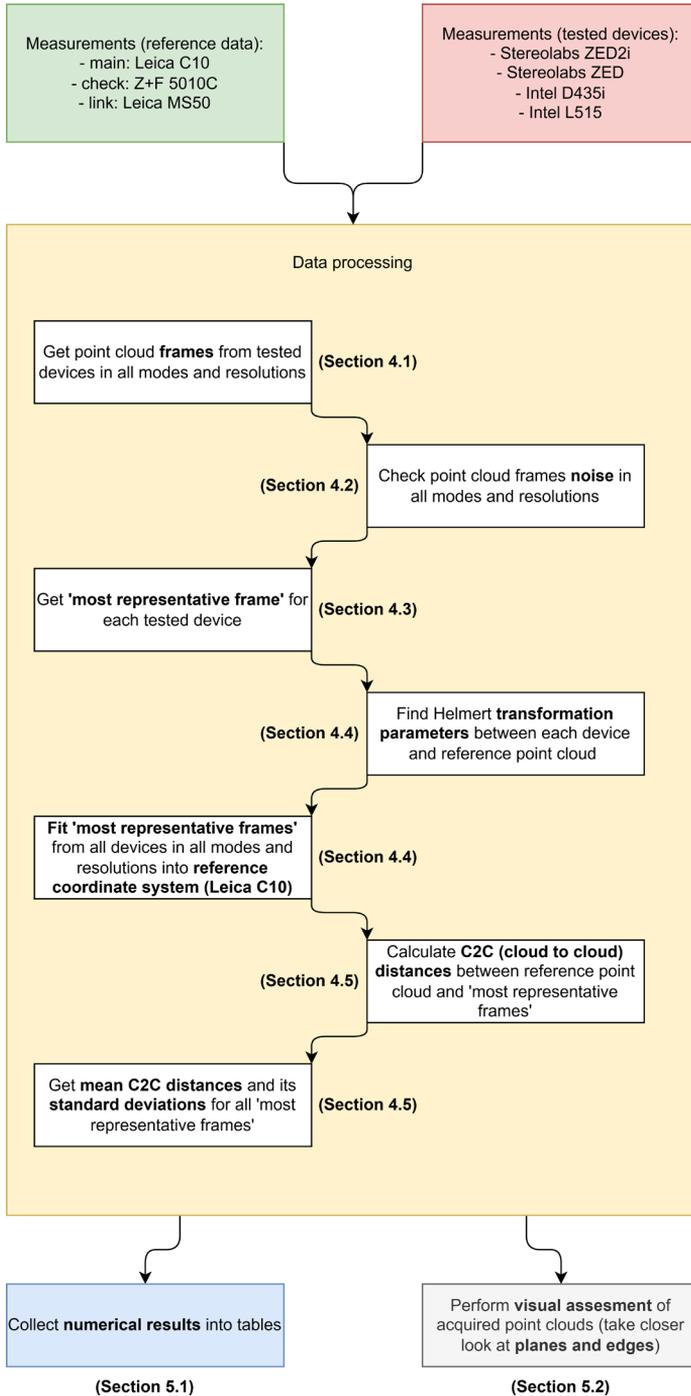
Measurements (reference data):
- main: Leica C10
- check: Z+F 5010C
- link: Leica MS50

Measurements (tested devices):
- Stereolabs ZED2i
- Stereolabs ZED
- Intel D435i
- Intel L515

Data processing

Get point cloud **frames** from tested devices in all modes and resolutions **(Section 4.1)**

(Section 4.2) Check point cloud frames **noise** in all modes and resolutions

Get **'most representative frame'** for each tested device **(Section 4.3)**

(Section 4.4) Find Helmert **transformation parameters** between each device and reference point cloud

Fit 'most representative frames' from all devices in all modes and resolutions into **reference coordinate system (Leica C10)** **(Section 4.4)**

(Section 4.5) Calculate **C2C (cloud to cloud) distances** between reference point cloud and 'most representative frames'

Get **mean C2C distances** and its **standard deviations** for all 'most representative frames' **(Section 4.5)**

Collect **numerical results** into tables

(Section 5.1)

Perform **visual assesment** of acquired point clouds (take closer look at **planes and edges**)

(Section 5.2)

Fig. 4. Research and data processing pipeline

4.1. Frame Selection

The Stereolabs ZED Depth Viewer software allowed for point-cloud generation from raw.svo files in four different modes: Performance, Quality, Ultra, and Neural. Each of these modes was characterized by different qualities of the point clouds that were generated, the use of computer resources, and, thus, the duration of the point-cloud-generation process. In Section 5, the quality of the results of each mode will be presented. The generated point clouds could be saved in the popular. ply format. The latest version of the ZED SDK 4.0 software was used for the study, with an "improved NEURAL depth mode, which offers even more accurate depth maps in challenging situations such as low-light environments and texture-less surfaces" [56].

For the Intel sensors, the export function to the.ply format in Intel RealSense Viewer v2.53.1 was used to save selected random frames from the point-cloud movie.

4.2. Point-Cloud Noise

In the first step of the data validation, the values of the point-cloud noise were checked; this was determined by comparing the cloud-to-cloud distances between raw sensor frames (without any modifications) in a given mode. After generating ten random but consecutive frames (i – diagram is shown in Figure 5) from each camera (S) at each resolution (R) and in each possible mode (M), a C2C (cloud-to-cloud) measurement analysis and standard statistical calculations were performed between each of them for this purpose (all possible pair combinations). CloudCompare point-cloud-processing software was used for this. It should be noted that the pairs being compared were also repetitive – once, a given cloud was the cloud being compared (A), and another time, it was the cloud to which the other clouds were compared (B); this was due to the small but present differences in the mean distances of A to B (as can be seen in Tables 3 and 5). Due to the large number of files that were generated, a Python script and the CloudComPy library [57] were used to calculate the cloud-to-cloud distances. The following parameters were computed:
– minimum distances between clouds,
– maximum distances between clouds,
– mean distances between clouds,
– variances of mean distances,
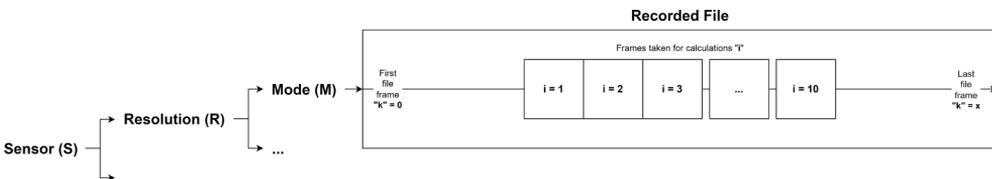– standard deviations of mean distances.

Fig. 5. Supplementary diagram on processing of survey data (point-cloud frames)

An extract of the results that were obtained for one camera is given in Table 3; it can be noticed that the "max" column reached values of up to several meters. This presents how noisy these sensors are, what their true nature is, and what maximum deviations the user can expect. Hence, raw and unfiltered data was used. Considering the impacts of the mentioned maximum values on the average and its standard deviation, they will have a negligible impact (outliers were below 0.01%); however, such distances did exist, so it is important to indicate this by presenting the real maximum distance.

Table 3. Sample part of cloud-to-cloud calculation results – measurement noise of given camera (in this case, S – Stereolabs ZED 2i camera, R – HD1080 resolution, M – Neural (n) mode, and $i_1...i_{10} = k_{40}...k_{49}$ – calculation frames start from 40th frame of recording)

Cloud B (S + R + M + k)	Cloud A (S + R + M + k)	max [m]	mean [m]	stdev [m]
zed2i HD1080 n40	zed2i HD1080 n41	1.595	0.036	0.051
zed2i HD1080 n40	zed2i HD1080 n42	3.467	0.035	0.052
...	...	...	...	...
zed2i HD1080 n40	zed2i HD1080 n49	3.952	0.049	0.078
zed2i HD1080 n41	zed2i HD1080 n40	2.638	0.036	0.050
zed2i HD1080 n41	zed2i HD1080 n42	4.022	0.035	0.058
...	...	...	...	...
zed2i HD1080 n41	zed2i HD1080 n49	3.777	0.048	0.084
zed2i HD1080 n42	zed2i HD1080 n40	1.330	0.034	0.046
zed2i HD1080 n42	zed2i HD1080 n41	1.434	0.034	0.053
...	...	...	...	...
zed2i HD1080 n42	zed2i HD1080 n49	2.844	0.040	0.053
...	...	...	...	...

Then, the following parameters were calculated separately for each resolution and mode from all possible pair combinations within the ten frames:
- mean distance d of compared clouds to each other, which was derived from measurement noise (mean mean),
- mean standard deviation of distance d (mean stdev),
- maximum mean distance d (max mean),
- maximum standard deviation of distance d (max stdev),
- median of mean distance d (median mean),
- median standard deviation of distance d (median stdev).

Example results are shown in Table 4 and Figure 6.

Table 4. Summary table of noise analysis for given camera
(here, Stereolabs ZED 2i camera at HD1080 resolution)

Mode	mean mean [m]	mean stdev [m]	max mean [m]	max stdev [m]	median mean [m]	median stdev [m]
Performance	0.077	0.174	0.113	0.386	0.074	0.122
Quality	0.035	0.046	0.050	0.071	0.034	0.045
Ultra	0.054	0.093	0.098	0.263	0.051	0.082
Neural	0.043	0.074	0.071	0.231	0.041	0.066
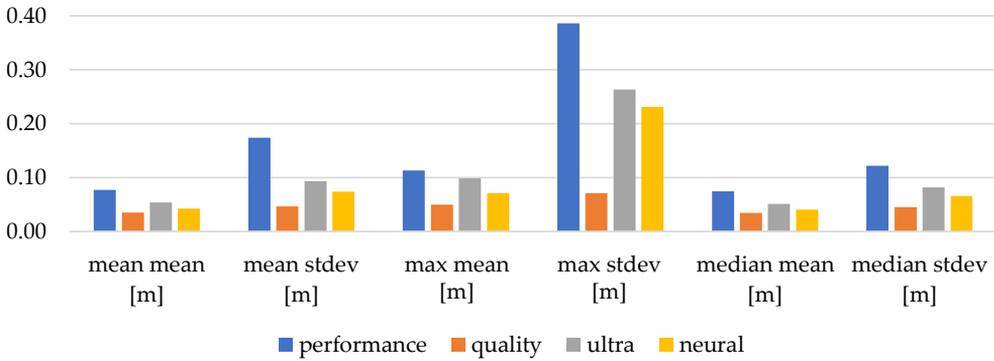
Fig. 6. Measurement noise for given camera, resolution, and mode
(here, Stereolabs ZED 2i camera at HD1080 resolution)

4.3. 'Most Representative Frame'

In each mode and resolution, the 'most representative frame' was selected for every ten pairs of frames. To do this, the mean distance results were first collated as a square matrix (Table 5); once, a given cloud was the cloud that was being compared (A), and another time, it was the cloud to which the other clouds were compared (B). The values of the mean distance relative to each A (vertical – SUM_B) and B (horizontal – SUM_A) point cloud were summed. Based on the values of these sums, their average was determined:

$$\text{MEAN OF SUM}_i = \frac{\text{SUM}_{Ai} + \text{SUM}_{Bi}}{2} \tag{1}$$

where $i \in \{1, 2, ..., 10\}$.

MEAN OF SUM represents the position that a given cloud had numerically in relation to the others. If the mean distances to all of the others were large, this meant that it was an outlier cloud; if the opposite was true (i.e., the average distances were small), the given cloud was close to the others. The cloud for which MEAN OF SUM had the smallest value was the cloud that was closest to all of the others and was, therefore, the 'most representative frame.' Examples of the mean-distance values for the ZED 2i camera (HD1080 resolution/Neural mode) are shown in Table 5.

Table 5. Summary of mean distances between clouds from consecutive frames (noise) for selected camera, with 'most representative frame' highlighted (Number 3) for Stereolabs ZED 2i camera (HD1080 resolution/Neural mode)

A / B	1	2	3	4	5	6	7	8	9	10	SUM$_A$ [m]	MEAN OF SUM [m]
1		0.036	0.035	0.063	0.046	0.032	0.037	0.046	0.037	0.049	0.381	0.399
2	0.036		0.035	0.060	0.043	0.042	0.033	0.038	0.040	0.048	0.375	0.393
3	0.034	0.034		0.050	0.037	0.032	0.031	0.036	0.036	0.040	0.332	0.349
4	0.061	0.057	0.045		0.039	0.046	0.049	0.042	0.057	0.036	0.432	0.479
5	0.053	0.048	0.045	0.060		0.037	0.034	0.033	0.041	0.035	0.386	0.365
6	0.037	0.048	0.036	0.054	0.036		0.033	0.039	0.038	0.037	0.357	0.350
7	0.045	0.041	0.041	0.071	0.036	0.035		0.039	0.032	0.043	0.385	0.355
8	0.050	0.041	0.039	0.051	0.030	0.038	0.034		0.041	0.034	0.357	0.357
9	0.045	0.049	0.045	0.070	0.044	0.042	0.034	0.048		0.048	0.425	0.396
10	0.057	0.055	0.045	0.047	0.033	0.038	0.041	0.036	0.045		0.398	0.384
SUM$_B$ [m]	0.418	0.411	0.365	0.527	0.343	0.343	0.325	0.358	0.368	0.371	–	–

This activity was repeated for all of the resolutions and their modes. For each set of modes at a given resolution, a comparison was also made (see Table 6 for an example). The frame with the lowest MEAN OF SUM from all of the resolutions and modes took part in determining the rotation (rotation matrix) of the clouds from a given sensor relative to the reference cloud.

For each camera, the resolution and mode for which the 'most representative frame' achieved the lowest noise value was selected. For the Stereolabs ZED 2i camera, this was HD1080 resolution and Quality mode (Table 6); interestingly, this was not the highest possible 2K resolution nor Neural mode (as was expected). The next step was to import this frame into CloudCompare. After loading four frames (one 'best' of the 'most representative frames' for each camera), a reference cloud (acquired by the Leica C10) was also added and cropped to a frame that the cameras

could cover with their fields of view. The reference scan from the Leica C10 (the cloud in the.pts format) was previously generated from the original Leica Cyclone software.

Table 6. MEAN OF SUM summary with 'most representative frames' highlighted for Stereolabs ZED 2i camera at HD1080 resolution

i	Neural [m]	Performance [m]	Quality [m]	Ultra [m]
1	0.399	0.664	0.293	0.458
2	0.393	0.646	0.296	0.487
3	0.349	0.715	0.317	0.471
4	0.479	0.772	**0.286**	0.486
5	0.365	0.670	0.314	0.489
6	0.350	0.768	0.358	0.489
7	0.355	0.716	0.288	0.455
8	0.357	0.659	0.312	0.608
9	0.396	0.640	0.330	0.455
10	0.384	0.716	0.360	0.455

4.4. Transformation to Reference Coordinate System

For each camera (and the best frame that was selected for it), a cloud-to-cloud alignment was applied based on five tie points that were selected after an initial rough transformation. These were measurement targets that were unambiguously identifiable on the point cloud (black and white targets, Leica HDS targets, and scanning spheres – visible in Figure 2). The resulting fit errors are summarized in Table 7. The Helmert transformation (6DOF) between each sensor and reference point cloud was estimated based on the measured tie points.

Table 7. Fit errors (RMS) of best 'most representative frame' point clouds of tested sensors to reference cloud from Leica C10

Camera	RMSE [m]
Stereolabs ZED 2i	0.027
Stereolabs ZED	0.395
Intel D435i	0.099
Intel L515	0.010

The derived 4 × 4 transformation matrix for each camera was separately applied to its 'most representative frame' at each resolution and in each mode, thus obtaining clouds that were fitted into the reference coordinate system.

4.5. Cloud-to-Cloud Distances

The next step was to recalculate the C2C distance (this time, relative to the cloud from the Leica C10). The results of the calculations are summarized below (Table 8).

Table 8. Summary of results of comparing point clouds from ZED 2i camera
against reference from Leica C10
(n – neural, p – performance, q – quality, u – ultra)

Reference	Compared	Max distance [m]	Mean distance [m]	Standard deviation [m]
Leica C10	zed2i HD2K n44	11.848	0.142	0.363
Leica C10	zed2i HD2K p43	12.753	0.279	0.539
Leica C10	zed2i HD2K q44	9.331	0.238	0.448
Leica C10	zed2i HD2K u45	11.832	0.216	0.444
Leica C10	zed2i HD720 n44	11.859	0.278	0.518
Leica C10	zed2i HD720 p42	25.298	0.797	2.223
Leica C10	zed2i HD720 q45	11.820	0.564	1.044
Leica C10	zed2i HD720 u44	11.859	0.278	0.518
Leica C10	zed2i HD1080 n42	12.103	0.135	0.372
Leica C10	zed2i HD1080 p48	19.457	0.293	0.657
Leica C10	zed2i HD1080 q43	12.568	0.212	0.409
Leica C10	zed2i HD1080 u46	11.810	0.192	0.450
Leica C10	zed2i VGA n43	13.335	0.256	0.527
Leica C10	zed2i VGA p40	25.083	1.030	2.410
Leica C10	zed2i VGA q43	8.560	0.432	0.613
Leica C10	zed2i VGA u49	20.080	0.514	1.399

Based on Table 8, a summary of the data on the mean distances of the acquired point clouds (relative to the reference clouds) was prepared. Table 9 shows these values in a color-coded scheme (from blue [smallest] to red [largest]). Similarly, the standard deviations are summarized in Table 10.

Table 9. Summary of mean distances of compared point clouds from ZED 2i camera relative to reference cloud from Leica C10
(with extreme values highlighted in color scheme mentioned above)

Resolution	Performance [m]	Quality [m]	Ultra [m]	Neural [m]
HD2K	0.279	0.238	0.216	0.142
HD1080	0.293	0.212	0.192	0.135
HD720	0.797	0.564	0.278	0.278
VGA	1.030	0.432	0.514	0.256

Table 10. Summary of standard deviations of mean distances of compared point clouds from ZED 2i camera relative to reference cloud from Leica C10
(with extreme values highlighted in color scheme mentioned above)

Resolution	Performance [m]	Quality [m]	Ultra [m]	Neural [m]
HD2K	0.539	0.448	0.444	0.363
HD1080	0.657	0.409	0.450	0.372
HD720	2.223	1.044	0.518	0.518
VGA	2.410	0.613	1.399	0.527

Since the calculated noise values significantly exceeded the manufacturer's declared values in the case of the Stereolabs ZED 2i camera, the obtained point clouds were additionally cut to the boundary of the test field due to artifacts that arose from the observation of the space outside the window. The distances and standard deviations were then recalculated based on the reduced point clouds. Detailed data is presented in Section 5.

5. Results

5.1. Sensors Precision and Accuracy

The data that was collected from all of the tested sensors was subjected to analysis according to the procedure that is described in Section 3 – Data Processing. To simplify the interpretation of the results, only the average values of "measurement noise" (Table 11) will be presented, along with the average distances between the clouds from the tested sensors and the reference cloud from the Leica C10 scanner (Table 12); these may serve as indicators that best reflect the quality of the measurement results of the tested devices. Both tables have also been color-coded on a scale from blue (lower values – better) to red (higher values – worse). Detailed results for each device at the individual calculation stages have been included in the supplementary data file.

Table 11. Measurement noise [m]
(average distances among ten consecutive measurement frames in all possible combinations)
for all devices (along with their corresponding resolutions and modes)

Mode	Stereolabs ZED 2i				
	Resolution				
	HD2K	HD1080	HD720	VGA	
Performance	0.095	0.077	0.200	0.232	
Quality	0.041	0.035	0.082	0.110	
Ultra	0.065	0.054	0.077	0.096	
Neural	0.048	0.043	0.077	0.111	

Mode	Stereolabs ZED				
	Resolution				
	HD2K	HD1080	HD720	VGA	
Performance	0.118	0.107	0.206	0.200	
Quality	0.090	0.083	0.174	0.195	
Ultra	0.105	0.090	0.196	0.193	
Neural	0.102	0.065	0.157	0.264	

Mode	Intel D435i					
	Resolution					
	1280 × 720	848 × 480	640 × 480	640 × 360	480 × 270	424 × 240
Default	0.031	0.033	0.024	0.032	0.034	0.033
Hand	0.029	0.029	0.026	0.032	0.033	0.035
High accuracy	0.022	0.024	0.019	0.026	0.028	0.029
High density	0.033	0.032	0.024	0.035	0.037	0.036
Medium density	0.028	0.031	0.022	0.030	0.030	0.029

Mode	Intel L515			
	Resolution			
	1024 × 768	640 × 480	320 × 240	
No ambient light	0.014	0.016	0.016	
Low ambient light	0.023	0.029	0.027	
Max range	0.014	0.016	0.017	
Short range	0.036	0.028	0.027	

Table 12. Average distances [m] among reference cloud from Leica C10
and 'most representative frames' from all devices
(along with their corresponding resolutions and modes)

Mode	Stereolabs ZED 2i				
	Resolution				
	HD2K	HD1080	HD720	VGA	
Performance	0.279	0.293	0.797	1.030	
Quality	0.238	0.212	0.564	0.432	
Ultra	0.216	0.192	0.278	0.514	
Neural	0.142	0.135	0.278	0.256	

Mode	Stereolabs ZED 2i – after segmentation				
	Resolution				
	HD2K	HD1080	HD720	VGA	
Performance	0.187	0.187	0.188	0.196	
Quality	0.157	0.158	0.161	0.217	
Ultra	0.128	0.122	0.131	0.175	
Neural	0.094	0.081	0.131	0.111	

Mode	Stereolabs ZED				
	Resolution				
	HD2K	HD1080	HD720	VGA	
Performance	2.483	2.354	2.556	2.422	
Quality	2.944	2.741	2.988	2.756	
Ultra	2.339	2.789	2.753	2.547	
Neural	2.671	2.758	2.932	3.002	

Mode	Intel D435i					
	Resolution					
	1280 × 720	848 × 480	640 × 480	640 × 360	480 × 270	424 × 240
Default	0.084	0.081	0.080	0.092	0.096	0.107
Hand	0.069	0.068	0.068	0.076	0.095	0.093
High accuracy	0.063	0.065	0.064	0.079	0.091	0.100
High density	0.078	0.074	0.071	0.081	0.091	0.102
Medium density	0.075	0.077	0.066	0.083	0.106	0.109

Mode	Intel L515		
	Resolution		
	1024 × 768	640 × 480	320 × 240
No ambient light	0.084	0.036	0.036
Low ambient light	0.012	0.012	0.010
Max range	0.046	0.039	0.028
Short range	0.012	0.011	0.011

5.2. Visual Assessment of Acquired Point Clouds

In the previous section, a detailed analysis of the point clouds that were obtained when using the examined sensors was carried out on the basis of the numerical values that characterized them. It should be noted, however, that these referred to the full imaged scenes without distinguishing any characteristic elements such as targets or planes (which are often the subject of observations). In addition, an important assessment of the observations of a rather qualitative nature can be made by making a visual comparison of the point clouds; in this way, the fidelity of the representation of the observed elements can be assessed.

In Figure 7, a reference point cloud that was obtained with the professional Leica C10 laser scanner is presented. Obviously, all of the elements (targets, spheres, floor, and walls) have been imaged correctly.

a)   b)
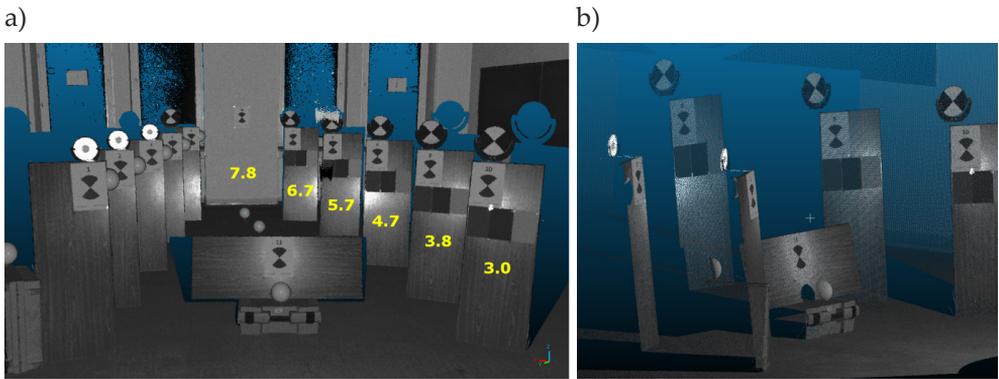
Fig. 7. Reference point cloud acquired by Leica C10 laser scanner: a) from scanner position (with distances in meters from scanner to planes analyzed); b) empty spaces between planes

In Figure 8, the same scene is shown as in Figure 7a; in this case, it was recorded by the other sensors from the same position as for the reference cloud. The weaker coverages for the L515 and D435i sensors are clearly visible; the sparse cloud for the former prevents a further comparative analysis for the planes, while only the two planes that are visible on the right (A, B) will be analyzed for the latter. On the other hand, both ZED devices were characterized by good ranges and fidelity of representation of the spatial elements. A total of five planes (labeled A–E in the figure) will be further analyzed.

Before performing a further analysis, it was useful to compare the clouds from the low-cost sensors with the reference cloud. In Figure 9, the compared clouds (in color scale) are shown against the reference cloud (in gray-scale). The discrepancies were most visible for the Stereolabs ZED and Intel D435i sensors, while the Intel L515 showed the best results (albeit, for a limited range).
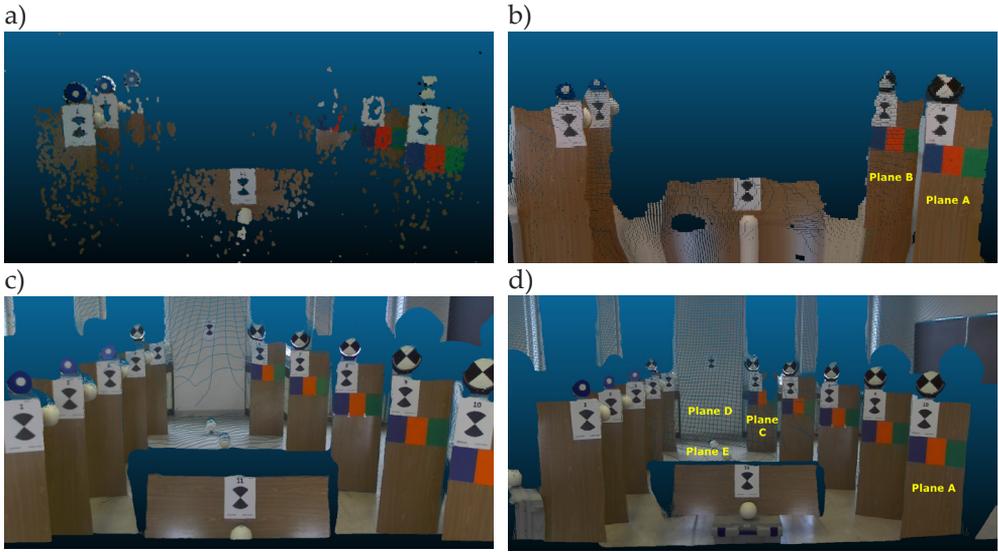
Fig. 8. Point clouds acquired by: a) L515; b) D435i; c) ZED; d) ZED 2i (with planes analyzed)
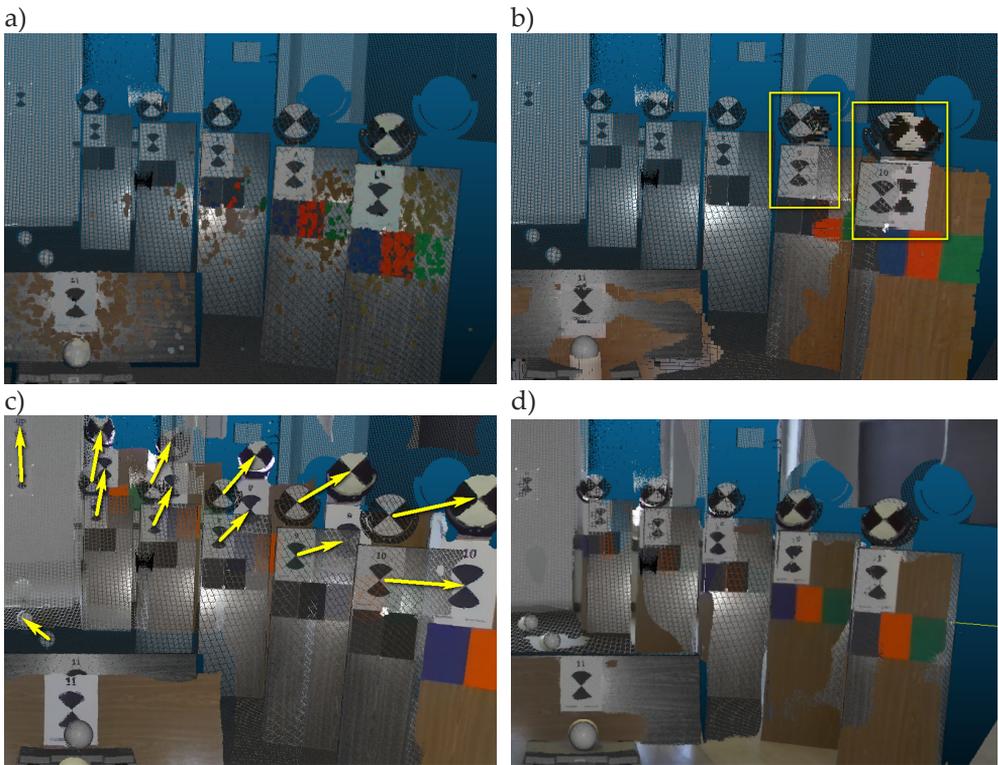

Fig. 9. Discrepancies between reference cloud and one obtained by following sensors: a) L515; b) D435i; c) ZED; d) ZED 2i (with clear differences marked in yellow)

For the correctness of distinguishing the separate elements of the observed scene, it is crucial to see what the spaces between the test field planes looked like; in fact, they were empty (as can be seen in the reference cloud – Fig. 7b). The same section of the test field for the other sensors is presented in Figure 10.
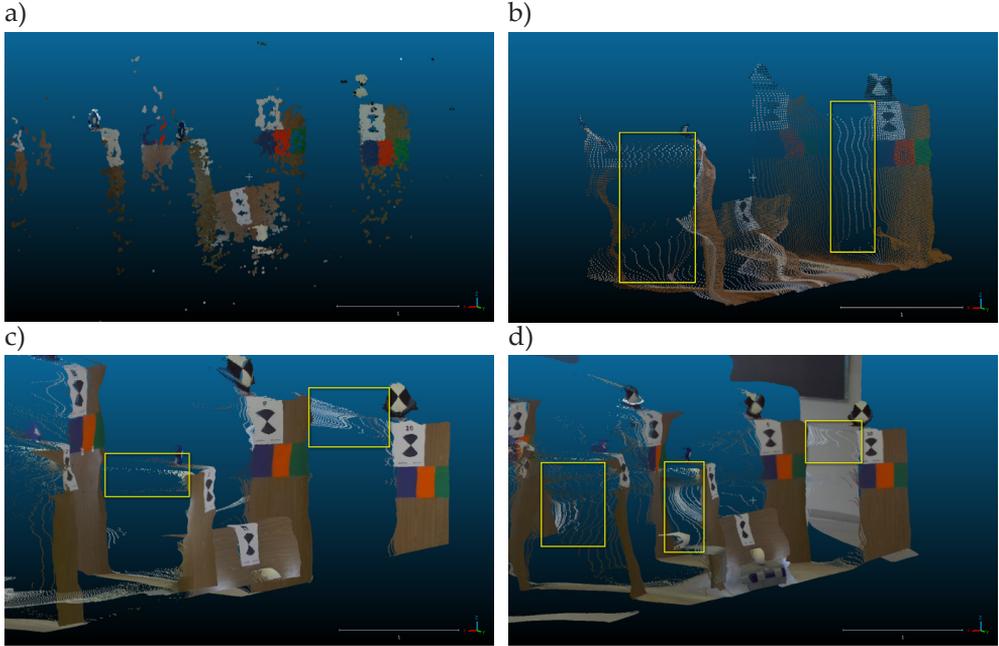
Fig. 10. Imaging empty spaces between planes for following sensors:
a) L515; b) D435i; c) ZED; d) ZED 2i
(with blurred edges marked in yellow)

Based on the tests, the data from the depth cameras was found to be significantly affected by artifacts that appeared in gaps between the planes (unlike from laser scanners). Groups of points appear on the edges of the imaged features, creating some non-existent surfaces (especially for the D435i sensor). These artifacts can be misinterpreted by systems that use depth cameras for determining the spatial positions of objects. In contrast, some single points appeared between the planes for the L515 sensor (which uses the LiDAR technology; this translated as measurement noise, is easy to filter out due to its random location. Admittedly, the imaging range was small, but this is a feature of this low-cost sensor. In order to determine the correctness of the plane representation by the various sensors, comparisons were made between selected planes and their counterparts from the reference data. Two sensors were excluded from the analyses: the L515 (due to its excessively sparse clouds and failure to reproduce the shapes of the planes), and the ZED (due to a probable problem with the scales of the imaged objects).

Figure 11 shows the comparison results for the D435i device; these were made on the planes within the imaging range of the sensor (i.e., A and B – Fig. 8b). Interestingly, an increase in the distance from the sensor to the plane did not indicate a decrease in the imaging accuracy.

a)

b)
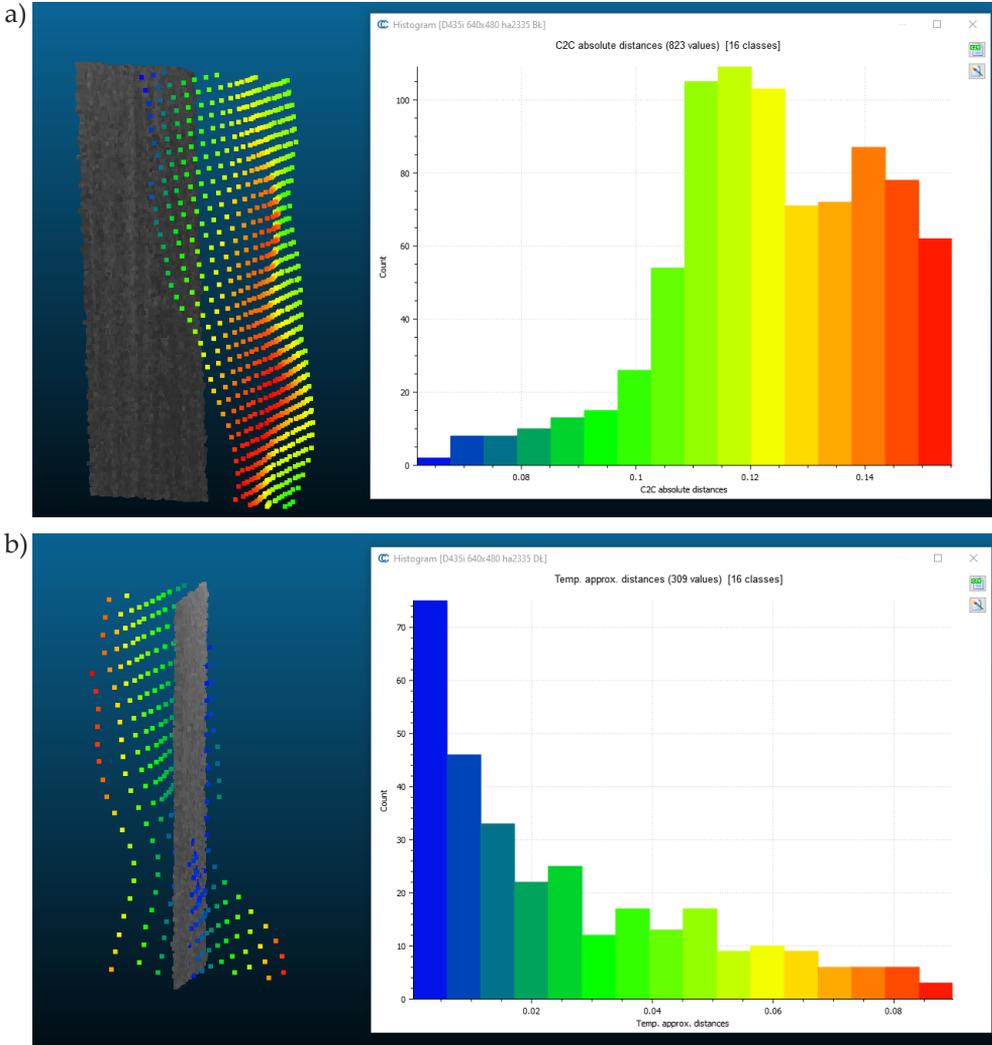
Fig. 11. Plane discrepancies that were recorded by D435i sensor relative to reference data for Plane A (a) and Plane B (b) (along with histograms of differences)

A similar comparison was made for the ZED 2i device (Fig. 12); however, this was done for a larger number of planes due to the larger imaging range (Fig. 9d). The obtained accuracy corresponded to the parameters that were stated by the manufacturer (i.e., 5% for the 3–15 m range).
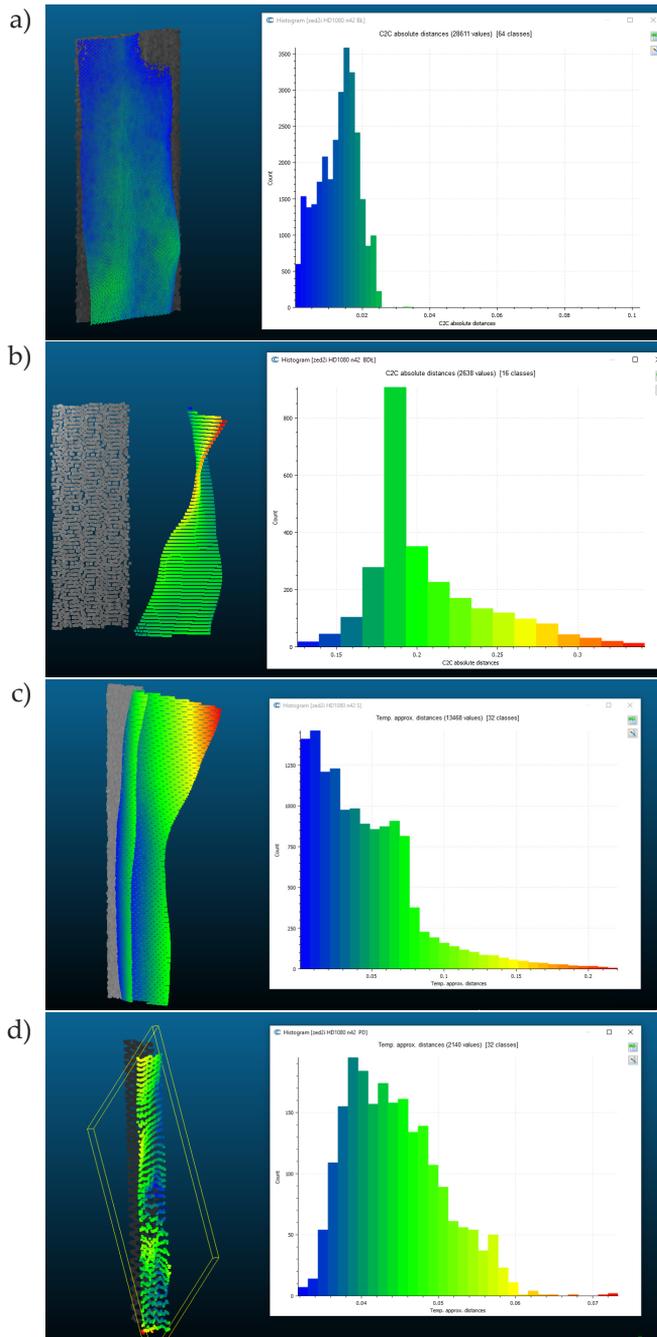
Fig. 12. Plane discrepancies that were recorded by ZED 2i sensor
relative to reference data for: a) Plane A; b) Plane C; c) Plane D; d) Plane E
(along with histograms of differences)

However, the impact of the distance on the accuracy cannot be clearly indicated, as a better result was obtained for the further Plan D (Fig. 12c) than for the closer Plane C (Fig. 12b). Interestingly, Plane E (representing the floor; i.e., the surface that was located unfavorably in relation to the device [Fig. 12d]) was imaged better than the vertical planes that were located at a similar distance.

6. Discussion

All of the tested devices allowed for acquisitions of 3D data from the closest surroundings (field of view). The Stereolabs ZED 2i device coped correctly (with an accuracy of ±8 cm) up to a range of about 10 m; above this value, however, the errors started to increase significantly.

The Stereolabs ZED unfortunately did not meet our expectations. At first glance, the results looked correct visually; as it turned out, however, they were not true-to-scale.

The Intel devices offered much shorter distances while being cheaper, less demanding, and more accurate. The Intel D435i offered good results (with an accuracy of ±6 cm), but the obtained cloud were quite sparse, and the working range was very short; however, they were quite suitable for detecting objects in their immediate surroundings (up to about 3 m).

On the other hand, the Intel L515 LiDAR offered an accuracy of ±1 cm – a surprisingly positive result in this price range. Unfortunately, the manufacturer's suggested range of 9 m was unattainable (even in complete darkness). In daylight, the real range was about 3–4 m, and in artificial light – about 4–5 m. In darkness, the achievable point-cloud density improved significantly, and the range increased to around 7 m.

Of all of the devices that were tested, the Intel sensors were the biggest positive surprises due to their price/accuracy ratios. On the contrary, the Stereolabs devices offered very good density of the generated point cloud, and they also increased their accuracy in Neural mode. Interestingly, the lowest noise between consecutive frames was obtained when operating in HD1080 resolution and Quality mode rather than in 2K and Neural mode (as was expected).

After the testing, it is possible to identify the resolutions and modes that performed best when working under standard conditions (with reasonable lighting and the availability of varied objects and environments):
– Stereolabs ZED 2i – HD1080 resolution, Neural mode;
– Stereolabs ZED – due to scaling of cloud, it is difficult to give clear answer, while HD1080 resolution and Neural mode had lowest noise factor;
– Intel D435i – 640 × 480 resolution (results that were comparable to maximum 1280 × 720, but with lower noise), High Accuracy mode;
– Intel L515 – 1024 × 768 resolution, but mode very dependent on current ambient light conditions.

The manufacturers' stated parameters for accuracy and maximum ranges should be taken with limited confidence; however, they were not far from reality (which speaks well of their quality).

When choosing the right device for a planned project, there are many factors to consider; the most important of these are as follows (the winning device for each category is provided in brackets):

- measurement distance (greatest distance – Stereolabs ZED 2i);
- lighting conditions (night/low light operation – Intel L515);
- computing capabilities of existing (or planned to be installed) central processing unit (no need for external unit – Intel sensors);
- power capabilities of system in which it is installed (Stereolabs devices needed least power due to lack of computing unit inside them);
- budget available (cheapest – Intel D435i [around $334]).

7. Conclusion

To summarize the obtained results (and having taken a number of factors into account), the overall test winner was the Stereolabs ZED 2i – provided it is running at HD1080 resolution and in Neural mode (as well as having adequate power and computing resources). However, the winner of the tests in terms of the accuracy of the obtained data was unquestionably the Intel L515; although still available for purchase, it has unfortunately been discontinued by the manufacturer.

Finally, the aim of the presented research (i.e., determining the actual accuracy of the depth sensors) was achieved. This may constitute the ground truth for assessing any accuracy losses that result from the operations of vehicles moving under actual conditions.

CRediT Author Contribution

All authors contributed to the study conception and design. All authors read and approved the final manuscript.

D. J.: conceptualization, methodology, software, validation, formal analysis, investigation, resources, data curation, writing – original draft preparation, writing – review and editing, visualization, project administration, funding acquisition.

Ł. O.: methodology, validation, formal analysis, investigation, resources, writing – review and editing, supervision.

P. K.: methodology, validation, formal analysis, investigation, resources, writing – original draft preparation, writing – review and editing.

Declaration of Competing Interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

The data (''Complete results.pdf'', ''raw.zip'', ''frames exported.zip'' and ''frames 'most average' transformed.zip'') that support the findings of this study are available on request from the corresponding author [D.J., janos@agh.edu.pl]. The data are not publicly available due to their high volume.

Use of Generative AI and AI-assisted Technologies

No generative AI or AI-assisted technologies were employed in the preparation of this manuscript.

References

[1] Raj T., Hashim F.H., Huddin A.B., Ibrahim M.F., Hussain A.: *A survey on LiDAR scanning mechanisms*. Electronics, vol. 9(5), 2020, 741. https://doi.org/10.3390/electronics9050741.

[2] Nam D.V., Gon-Woo K.: *Solid-state LiDAR based-SLAM: A concise review and application*, [in:] *2021 IEEE International Conference on Big Data and Smart Computing (BigComp), Jeju Island, Korea (South), 2021*, IEEE, Piscataway 2021, pp. 302–305. https://doi.org/10.1109/BigComp51126.2021.00064.

[3] Cabo C., Del Pozo S., Rodríguez-Gonzálvez P., Ordóñez C., González-Aguilera D.: *Comparing Terrestrial Laser Scanning (TLS) and Wearable Laser Scanning (WLS) for individual tree modeling at plot level*. Remote Sensing, vol. 10(4), 2018, 540. https://doi.org/10.3390/rs10040540.

[4] Wei W., Shirinzadeh B., Nowell R., Ghafarian M., Ammar M.M.A., Shen T.: *Enhancing solid state LiDAR mapping with a 2D spinning LiDAR in urban scenario SLAM on ground vehicles*. Sensors, vol. 21(5), 2021, 1773. https://doi.org/10.3390/s21051773.

[5] Benedek C., Majdik A., Nagy B., Rozsa Z., Sziranyi T.: *Positioning and perception in LIDAR point clouds*. Digital Signal Processing, vol. 119, 2021, 103193. https://doi.org/10.1016/j.dsp.2021.103193.

[6] Kutila M., Pyykönen P., Holzhüter H., Colomb M., Duthon P.: *Automotive LiDAR performance verification in fog and rain*, [in:] *2018 21st International Conference on Intelligent Transportation Systems (ITSC), Maui, HI, USA, 2018*, IEEE, Piscataway 2018, pp. 1695–1701. https://doi.org/10.1109/ITSC.2018.8569624.

[7] Li K., Li M., Hanebeck U.D.: *Towards high-performance solid-state-LiDAR-inertial odometry and mapping*. IEEE Robotics and Automation Letters, vol. 6(3), 2021, pp. 5167–5174. https://doi.org/10.1109/lra.2021.3070251.

[8] Nehme H., Aubry C., Solatges T., Savatier X., Rossi R., Boutteau R.: *LiDAR-based structure tracking for agricultural robots: Application to autonomous navigation in vineyards*. Journal of Intelligent & Robotic Systems, vol. 103, 2021, 61. https://doi.org/10.1007/s10846-021-01519-7.

[9] Haddeler G., Aybakan A., Akay M.C., Temeltas H.: *Evaluation of 3D LiDAR sensor setup for heterogeneous robot team*. Journal of Intelligent & Robotic Systems, vol. 100, 2020, pp. 689–709. https://doi.org/10.1007/s10846-020-01207-y.

[10] Murtiyoso A., Grussenmeyer P., Landes T., Macher H.: *First assessments into the use of commercial-grade solid state lidar for low cost heritage documentation*. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XLIII-B2-2021, 2021, pp. 599–604. https://doi.org/10.5194/isprs-archives-xliii-b2-2021-599-2021.

[11] Boehler W., Vicent M.B., Marbs A.: *Investigating laser scanner accuracy*. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XXXIV-5/C15 [CIPA 2003 XVIII International Symposium], 2003, pp. 696–701.

[12] Lee K.H., Ehsani R.: *Comparison of two 2D laser scanners for sensing object distances, shapes, and surface patterns*. Computers and Electronics in Agriculture, vol. 60(2), 2008, pp. 250–262. https://doi.org/10.1016/j.compag.2007.08.007.

[13] Jafri S.R.U.N., Shamim S., Faraz S.M., Ahmed A., Yasir S.M., Iqbal J.: *Characterization and calibration of multiple 2D laser scanners*. PLOS ONE, vol. 17(7), 2022, e0272063. https://doi.org/10.1371/journal.pone.0272063.

[14] Aijazi A.K., Malaterre L., Trassoudaine L., Checchin P.: *Systematic evaluation and characterization of 3D solid state LiDAR sensors for autonomous ground vehicles*. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XLIII-B1-2020, 2020, pp. 199–203. https://doi.org/10.5194/isprs-archives-xliii-b1-2020-199-2020.

[15] Jozkow G., Wieczorek P., Karpina M., Walicka A., Borkowski A.: *Performance evaluation of sUAS equipped with Velodyne HDL-32E LiDAR sensor*. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XLII-2/W6, 2017, pp. 171–177. https://doi.org/10.5194/isprs-archives-xlii-2-w6-171-2017.

[16] Lam T.F., Blum H., Siegwart R., Gawel A.: *SL Sensor: An open-source, real-time and robot operating system-based structured light sensor for high accuracy construction robotic applications*. Automation in Construction, vol. 142, 2022, 104424. https://doi.org/10.1016/j.autcon.2022.104424.

[17] Plaß B., Emrich J., Götz S., Kernstock D., Luther C., Klauer T.: *Evaluation of point cloud data acquisition techniques for scan-to-BIM workflows in healthcare*, [in:] *FIG e-Working Week 2021: Smart Surveyors for Land and Water Management – Challenges in a New Reality, the Netherlands, 21–25 June 2021*, FIG, 2021.

[18] Breitbarth A., Hake C.B., Notni G.: *Measurement accuracy and practical assessment of the lidar camera Intel RealSense L515*, [in:] Lehmann P., Osten W.,

Albertazzi Gonçalves Jr. A. (eds.), *Optical Measurement Systems for Industrial Inspection XII*, Proceedings of SPIE, vol. 11782, SPIE – International Society for Optics and Photonics, 2021, pp. 218–229. https://doi.org/10.1117/12.2592570.

[19] Zollhöfer M.: *Commodity RGB-D sensors: Data acquisition*, [in:] Rosin P., Lai Y.K., Shao L., Liu Y. (eds.), *RGB-D Image Analysis and Processing*, Advances in Computer Vision and Pattern Recognition, Springer, Cham 2019, pp. 3–13. https://doi.org/10.1007/978-3-030-28603-3_1.

[20] Rodríguez-Gonzálvez P., Guidi G.: *RGB-D sensors data quality assessment and improvement for advanced applications*, [in:] Rosin P., Lai Y.K., Shao L., Liu Y. (eds.), *RGB-D Image Analysis and Processing*, Advances in Computer Vision and Pattern Recognition, Springer, Cham 2019, pp. 67–86. https://doi.org/10.1007/978-3-030-28603-3_4.

[21] Carfagni M., Furferi R., Governi L., Santarelli C., Servi M., Uccheddu F., Volpe Y.: *Metrological and critical characterization of the Intel D415 stereo depth camera*. Sensors, vol. 19(3), 2019, 489. https://doi.org/10.3390/s19030489.

[22] Moghari M.D., Noonan P., Henry D.L., Fulton R., Young N., Moore K., Kyme A.: *Characterization of the Intel RealSense D415 stereo depth camera for motion-corrected CT imaging*, [in:] *2019 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), Manchester, UK, 2019*, IEEE, Piscataway 2019, pp. 1–3. https://doi.org/10.1109/NSS/MIC42101.2019.9059935.

[23] Curto E., Araujo H.: *An experimental assessment of depth estimation in transparent and translucent scenes for Intel RealSense D415, SR305 and L515*. Sensors, vol. 22(19), 2022, 7378. https://doi.org/10.3390/s22197378.

[24] Lourenço F., Araujo H.: *Intel RealSense SR305, D415 and L515: experimental evaluation and comparison of depth estimation*, [in:] Farinella G.M., Radeva P., Braz J., Bouatouch K. (eds.), *Proceedings of the 16th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, Volume 4: VISAPP*, SciTePress, Setúbal 2021, pp. 362–369. https://doi.org/10.5220/0010254203620369.

[25] Servi M., Mussi E., Profili A., Furferi R., Volpe Y., Governi L., Buonamici F.: *Metrological characterization and comparison of D415, D455, L515 RealSense devices in the close range*. Sensors, vol. 21(22), 2021, 7770. https://doi.org/10.3390/s21227770.

[26] Tadic V.: *Study on automatic electric vehicle charging socket detection using ZED 2i depth sensor*. Electronics, vol. 12(4), 2023, 912. https://doi.org/10.3390/electronics12040912.

[27] Connolly L., O'Gorman D., Garland J., Tobin E.F.: *Spatial mapping of light aircraft with stereo-vision camera for use on unmanned aircraft system for defect localisation*, [in:] *2023 International Conference on Unmanned Aircraft Systems (ICUAS), Warsaw, Poland*, IEEE, Piscataway 2023, pp. 413–418. https://doi.org/10.1109/icuas57906.2023.10155987.

[28] Ortiz L.E., Cabrera V.E., Gonçalves L.M.: *Depth data error modeling of the ZED 3D vision sensor from stereolabs*. Electronic Letters on Computer Vision and Image Analysis, vol. 17(1), 2018, pp. 1–15. https://doi.org/10.5565/rev/elcvia.1084.

[29] Skoczeń M., Ochman M., Spyra K., Nikodem M., Krata D., Panek M., Pawłowski A.: *Obstacle detection system for agricultural mobile robot application using RGB-D cameras*. Sensors, vol. 21(16), 2021, 5292. https://doi.org/10.3390/s21165292.

[30] Xue G., Li R., Liu S., Wei J.: *Research on underground coal mine map construction method based on LeGO-LOAM improved algorithm*. Energies, vol. 15(17), 2022, 6256. https://doi.org/10.3390/en15176256.

[31] Castorena J., Puskorius G.V., Pandey G.: *Motion guided LiDAR camera self-calibration and accelerated depth upsampling for autonomous vehicles*. Journal of Intelligent & Robotic Systems, vol. 100, 2020, pp. 1129–1138. https://doi.org/10.1007/s10846-020-01233-w.

[32] Liu J., Zhan X., Chi C., Zhang X., Zhai C.: *Robust extrinsic self-calibration of camera and solid state LiDAR*. Journal of Intelligent & Robotic Systems, vol. 109, 2023, 81. https://doi.org/10.1007/s10846-023-02015-w.

[33] Słomiński S., Sobaszek M.: *Intelligent object shape and position identification for needs of dynamic luminance shaping in object floodlighting and projection mapping*. Energies, vol. 13(23), 2020, 6442. https://doi.org/10.3390/en13236442.

[34] Liu L.S., Lin J.F., Yao J.X., He D.W., Zheng J.S., Huang J., Shi P.: *Path planning for smart car based on Dijkstra algorithm and dynamic window approach*. Wireless Communications and Mobile Computing, vol. 2021, 2021, 8881684. https://doi.org/10.1155/2021/8881684.

[35] Chghaf M., Rodriguez S., El Ouardi A.: *Camera, LiDAR and multi-modal SLAM systems for autonomous ground vehicles: A survey*. Journal of Intelligent & Robotic Systems, vol. 105, 2022, 2. https://doi.org/10.1007/s10846-022-01582-8.

[36] Dworak D., Baranowski J.: *Adaptation of Grad-CAM method to neural network architecture for LiDAR pointcloud object detection*. Energies, vol. 15(13), 2022, 4681. https://doi.org/10.3390/en15134681.

[37] Holder M., Elster L., Winner H.: *Digitalize the twin: A method for calibration of reference data for transfer real-world test drives into simulation*. Energies, vol. 15(3), 2022, 989. https://doi.org/10.3390/en15030989.

[38] Tadic V., Toth A., Vizvari Z., Klincsik M., Sari Z., Sarcevic P., Sarosi J., Biro I.: *Perspectives of RealSense and ZED depth sensors for robotic vision applications*. Machines, vol. 10(3), 2022, 183. https://doi.org/10.3390/machines10030183.

[39] Intel RealSense: Intel® RealSense™ LiDAR Camera L515, Datasheet, Rev 003, 2021.

[40] Stereolabs: ZED 2i Camera and SDK Overview, Rev 1, 2022.

[41] Stereolabs: ZED 2i. https://www.stereolabs.com/zed-2i/ [access: 18.08.2023].

[42] Stereolabs: ZED Camera and SDK Overview, Rev 2, 2018.

[43] Intel RealSense: Intel® RealSense™ Product Family D400 Series, Datasheet, Rev 015, 2023.

[44] Intel RealSense: Intel® RealSense™ Depth Camera D435i. https://www.intelrealsense.com/depth-camera-d435i/ [access: 18.08.2023].

[45] Intel RealSense: Intel® RealSense™ LiDAR Camera L515. https://www.intelrealsense.com/lidar-camera-l515/ [access: 18.08.2023].

[46] Cabrera E.V., Ortiz L.E., Silva B.M.D., Clua E.W., Gonçalves L.M.: *A versatile method for depth data error estimation in RGB-D sensors*. Sensors, vol. 18(9), 2018, 3122. https://doi.org/10.3390/s18093122.

[47] Zhang L., Xia H., Qiao Y.: *Texture synthesis repair of RealSense D435i depth images with object-oriented RGB image segmentation*. Sensors, vol. 20(23), 2020, 6725. https://doi.org/10.3390/s20236725.

[48] Trosin M., Dekemati I., Szabó I.: *Measuring soil surface roughness with the RealSense D435i*. Acta Polytechnica Hungarica, vol. 18(6), 2021, pp. 141–155.

[49] Forsey-Smerek A., Paige C., Ward F., Haddad D.D., Sanneman L., Todd J., Heldmann J., Lim D., Newman D.: *Assessment of depth data acquisition methods for virtual reality mission operations support tools*, [in:] *2022 IEEE Aerospace Conference (AERO), Big Sky, MT, USA, 2022*, IEEE, Piscataway 2022, pp. 1–14. https://doi.org/10.1109/AERO53065.2022.9843571.

[50] Intel RealSense: Intel® RealSense™ LiDAR Camera L515, Section Frequently Asked Questions. https://www.intelrealsense.com/lidar-camera-l515/ [access: 18.08.2023].

[51] Abbas M.A., Setan H., Majid Z., Chong A.K., Idris K.M., Aspuri A.: *Calibration and accuracy assessment of Leica ScanStation C10 terrestrial laser scanner*, [in:] Abdul Rahman A., Boguslawski P., Gold C., Said M. (eds.), *Developments in Multidimensional Spatial Data Models*, Lecture Notes in Geoinformation and Cartography, Springer, Berlin, Heidelberg 2013, pp. 33–47. https://doi.org/10.1007/978-3-642-36379-5_3.

[52] Antanavičiūtė U., Obuchovski R., Paršeliūnas E.K., Popovas M.G.D., Šlikas D.: *Some issues regarding the calibration of the terrestrial laser scanner Leica Scanstation C10*. Geodesy and Cartography, vol. 39(3), 2013, pp. 138–143. https://doi.org/10.3846/20296991.2013.840356.

[53] Leica Geosystems AG: Leica ScanStation C10, Product Specification, Heerbrugg, Switzerland, 2011.

[54] Zoller+Fröhlich: Z+F IMAGER® 5010 C, Datasheet, 2013.

[55] Leica Geosystems AG: Leica Nova MS50, Datasheet, Heerbrugg, Switzerland, 2013.

[56] Stereolabs: ZED SDK 4.0, Release Notes. https://www.stereolabs.com/developers/release/ [access: 18.08.2023].

[57] OpenFields: CloudComPy. https://www.simulation.openfields.fr/index.php/projects/cloudcompy [access: 18.08.2023].