

Article

Metrological and Critical Characterization of the Intel D415 Stereo Depth Camera

Monica Carfagni¹, Rocco Furferi^{1*}, Lapo Governi, Chiara Santarelli¹, Michaela Servi¹, Francesca Uccheddu¹ and Yary Volpe¹

¹ Department of Industrial Engineering, University of Florence, 50139 Firenze, Italy (e-mail: monica.carfagni@unifi.it; rocco.furferi@unifi.it; lapo.governi@unifi.it; chiara.santarelli@unifi.it; michaela.servi@unifi.it; francesca.uccheddu@unifi.it; yary.volpe@unifi.it)

* Correspondence: rocco.furferi@unifi.it

Abstract: Low-cost RGB-D cameras are increasingly used in several research fields including human-machine interaction, safety, robotics, biomedical engineering and even Reverse Engineering applications. Among the plethora of commercial devices, the Intel RealSense cameras proved to be among the best suitable devices, providing a good compromise between cost, ease of use, compactness and precision. Released on the market in January 2018, the new Intel model RealSense D415 has a wide acquisition range (i.e. ~160-10000 mm) and a narrow field of view to capture objects in rapid motion. Given the unexplored potential of this new device, especially when used as a 3D scanner, the present work aims to characterize and to provide metrological considerations on the RealSense D415. In particular, tests are carried out to assess the device performances in the near range (i.e. 100-1000 mm). Characterization is performed by integrating the guidelines of the existing standard (i.e. the German VDI/VDE 2634 part 2 normative) with a number of literature-based strategies. Performance analysis is finally compared against latest close-range sensors, thus providing a useful guidance for researchers and practitioners aiming to use RGB-D cameras in Reverse Engineering applications.

Keywords: Reverse Engineering, RealSense D415, depth camera, device characterization, VDI/VDE normative, active stereo, performance comparison

1. Introduction and Background

3D optical systems have found popularity among numerous fields of application spanning from robotics [1], automotive [2], industry [3] and mechanical engineering, to biomedical field [4][5,6]. Their success is mainly due to the recent developments which have allowed the creation of devices that are less expensive, yet accurate and compact.

Microsoft's original Kinect hardware was powered in September 2009 by PrimeSense [7]. The Israeli company pioneered the technology of projecting an infrared dots pattern onto a scene and detecting them with an IR camera to assess depth information. The output of the Kinect in its first version, was a 320 x 240 depth map with 2,048 levels of depth values, based on the projected IR speckle pattern. Later, other companies released low cost Kinect like depth cameras (i.e. Asus Xtion [8], Astra Pro 3D [9], Occipital Structure Sensor [10]); such sensors as solely human natural interface tools, were widely regarded as a flop for gaming, nevertheless the revolutionary depth-sensing tech ended up being a huge boost for robotics and machine vision [11].

In 2013, Apple bought PrimeSense and the depth cameras technology continued to evolve; Kinect 2.0 for the Xbox One replaced the PrimeSense technology with Microsoft's own time of flight technology [12] resulting in much higher accuracy and resolution. In 2016, Lenovo launched the Phab 2 Pro, the first smartphone to implement Google's Tango technology for augmented reality and machine vision, which is also based on infrared depth detection [13]. In late 2017, Apple released the iPhone X including Kinect like miniature depth-sensor [14]. Unlike the original Kinect, which was built to track motion in a whole living room, the sensor is primarily designed for scanning faces and powers Apple's Face ID feature [15].

Meanwhile, Intel also built its own depth sensors family, Intel RealSense [16], and in 2015 worked with Microsoft to power Windows Hello a 3D face recognition way to sign in to Windows 10 devices [17].

Despite the fact that Intel RealSense devices have only appeared on the market in recent years, they have been adopted in several fields. Among the many applications we find posture and gesture interaction systems and human interaction design [18–22], Interactive AI toy for children [23], security and robotics [24–26], and also medical and human care [27–29].

RealSense technology basically consists of vision processors, depth and tracking modules and depth cameras, supported by an open source multi-platform SDK called *librealsense* [30] that simplifies camera support for software developers and third-party system integrators. To overcome previous camera releases and strengthen their leading position in the market, Intel launched two new depth cameras in January 2018: the D415 and the D435 models. Such devices differ each other mainly in the field of view (FOV) angles and in the exposition time of the camera integrated shutter. The larger FOV of the Intel RealSense D435 depth camera translates into a minimization of blind spots, thus making it better for use cases such as robotics; the global shutter provides better performance when capturing high-speed movements, avoiding depth image blurring or shooting in low-light situations. Having a smaller FOV, the Intel RealSense D415 has a higher pixel density, thus resulting in higher resolution. Thereby, when accuracy is paramount (e.g. for 3D scanning applications), the Intel RealSense D415 promises to provide better results especially when used in the short range (i.e. <1 m).

Since the use of RGB-D cameras as a low-cost 3D scanner spread out in many applications, a comprehensive characterization of this new device is needed to tune up the best device and the related parameters settings for each scanning scenario. In a recent work, Giancola et al [31] have proposed a characterization for Microsoft Kinect V2, Orbbec Astra S and the Intel D400 series. For each of these devices two types of experiments were performed, one for the pixel-wise and the other for sensor-wise characterization to evaluate, respectively, the accuracy of the cameras at different distances and the quality in the reconstruction of known geometries. To evaluate the quality of the reconstruction of known geometries (i.e. sensor-wise characterization), the distribution of the distance between the acquired points and the actual geometric models was measured; during these tests, known planes, cylinders and spheres were acquired. The pixel-wise characterization of both Microsoft Kinect V2 and Orbbec Astra S camera, was done by placing the cameras on a photographic tripod and aligning it with a white planar target mounted on an anthropomorphic robot with an arm reach of 1200 mm and repeating the test four times to cover the distance from 800 mm to 4200 mm. As for the pixel-wise characterization of the Intel D400 cameras, the setup involved the use of two coordinated robots. The camera is fixed on one of the two robots and the target on the other thus allowing the entire range cover in a single test (e.g. without moving the setup). Author in [31] reported as expected, that the best depth camera performance (in terms of pixel wise accuracy and reconstruction of known geometries accuracy) are in the closest working range.

To pave the way towards a global performance measurement standardisation, thus providing the users with a comprehensive analysis of the camera limits and strengths in the best case of a close-range 3D scanner application scenario, additional tests are needed. Unfortunately, nowadays the international community has not yet released a recognized standard for the non-contact 3D imaging systems that includes depth camera systems.

The German standard VDI/VDE 2634 Part 2 (“Optical 3-D measuring systems”) [32] is in fact, the only recognised standard providing a number of guidelines for the acceptance and verification of optical 3-D measuring systems. It consists on a set of acceptance tests mainly based on evaluating three quality parameters: probing error, sphere spacing error, and flatness. Measurements are carried out using prescribed artefacts including a sphere, ball bar, and flat object. The guidelines indicate the exact size of the known artefacts according to the acquisition volume; in particular, the artefacts size is related to the length L_0 of the diagonal of the 3D imager’s conic field of view (as explained in Section 3). For each of these objects, both an acceptance test (i.e. calibrated artefacts are measured) and a re-verification test (i.e. the test is repeated over time) are uniquely defined. The tests are accepted if the error lies inside the limits specified by the manufacturer. Three errors are introduced by the standard:

1) the Probing Error P which describes the characteristic error of the system within a small part of the measurement volume; 2) the Sphere Spacing Error SS which demonstrates the ability of the system to measure lengths; 3) Flatness Measurement Error F i.e. the range of the signed distances of the measured points from the best-fit plane. Based on such a normative, in a previous work [33], the authors proposed a metrological and critical characterization for the previous RealSense RGB-D camera, the Intel SR300. The proposed methodology for the sensor characterization is delivered by integrating VDI/VDE guidelines with the recent characterization strategies provided in scientific literature [31,34–36].

Inspired by such a work, which allowed a full description of metrological properties of SR300, the main aim of this paper is to characterize and to provide metrological considerations on the Intel RealSense D415 depth sensor taking into account both current professional normative and best literature practices.

Additionally, the performance of D415 is compared against the RealSense SR300 and other latest short-range devices (i.e. Primesense Carmine 1.09 and Kinect v2), thus feeding the current critical discussion on this category of devices. This will allow researchers and practitioners' choice the optimal device for their own reverse engineering application.

The paper is organized as follows. Section 2 presents the Intel RealSense D415 depth camera specifications and working principle. The devised test set is presented in Section 3. Finally, experiments are discussed in Section 4 and conclusions are drafted.

2. Intel RealSense D415 Depth Camera

The Intel RealSense Depth Camera has been designed to equip devices with the ability to see, understand, interact with, and learn from their environment. The D415 features the Intel RealSense Vision D4 processor with high-resolution depth (up to 1280x720 at 30 frames per second), long-range capability (up to about 10 m), rolling shutter technology and, as said, a narrow field of view ideal for precise measurements. RealSense Vision D4 is a vision processor based on 28 nanometer (nm) process technology for real-time calculation of stereo depth data.

The device has a very compact depth camera (dimensions: 99mm x 20mm x 23mm, weight: 72 gr), that can either be integrated in computers and mobile devices or used as a self-standing device. Moreover, it comes with a color camera and a depth camera system, made by two IR cameras, and by an IR projector (Figure 1).

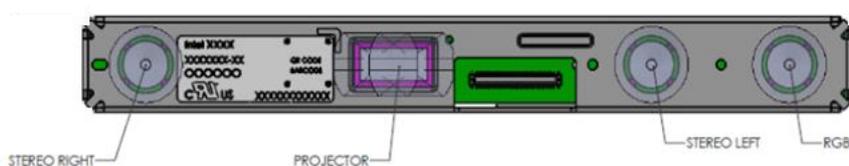


Figure 1. Components location inside the RealSense D415 camera

The infrared projector improves the ability of the stereo camera system to determine depth by projecting a static infrared pattern (Figure 2) on the scene to increase texture on low texture scenes. The left and right imagers capture the scene and sends raw image data to the Vision Processor, which calculates depth values for each pixel in the image by correlating points on the left to the right image. The depth pixel values are processed to generate a depth frame. This active stereo depth computation makes this depth camera suitable for acquisitions both indoors and outdoors under reasonable illumination. To this purpose the IR projector intensity can be tuned manually according to the environment lighting conditions. The depth data generated with stereo vision uses the left image as the reference for stereo matching resulting in a non-overlapping region on the field of view of the left and right image, therefore we will not have depth data at the left edge of the frame (Figure 2 in red). Closer scenes result in a wider invalid depth band than scenes at further distances.

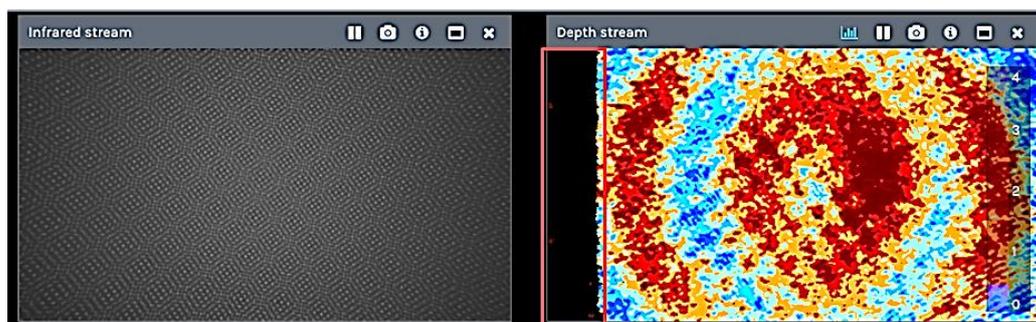


Figure 2. Static infrared dot pattern projected on a wall (left); In red is highlighted the non-overlapping region on the field of view of the left and right image (right)

The SR300 and the new D415 have different working principles for depth measurement (Active stereo vs. Structured light), which results in different performance as depicted in Figure 3. The figure shows a qualitative comparison of the two devices in representing a flat surface acquired at different camera distances: in grey are depicted the point clouds of a plane acquired at increasing distance using the SR300 model; the plane acquired with the D415 camera are instead described by the colored point clouds. It can be observed that the new device has a considerably lower error both in terms of planarity and of distance accuracy.

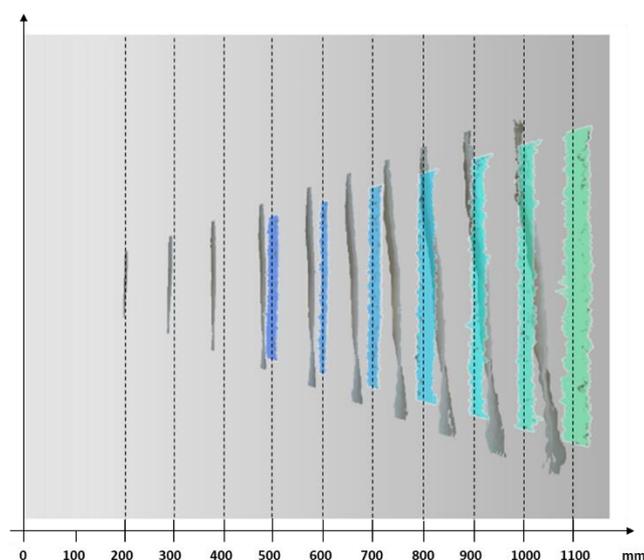


Figure 3. Qualitative evaluation of the error introduced in measuring a flat surface, and comparison between SR300 (grey point clouds) and D415 (colored point clouds) models.

The D415 has a narrow field of view solution using rolling shutter sensors; this means that very fast-moving objects or fast movement of the device (such as quick horizontal panning) can result in slightly distorted images. The Intel RealSense D400 series is supported with cross-platform and open source Intel RealSense SDK 2.0 [30], a library that allows to configure the camera with several internal settings (both for acquisition and post-processing). Furthermore, a set of visual pre-sets to improve performances are available. Such versatility allows users to arrange the best possible setup for the task under investigation.

The D415 is equipped with a color camera with resolution up to 1920x1080 provides texture information to be superposed on the depth data; the spatial resolution of the structured light depth camera of the Intel RealSense D415 is up to HD (1280 x 720), in a working depth range declared by the manufacturer equal to ~160-10000 mm. The camera works at different pixel resolutions corresponding to different minimum depth values, point density and framed area. The specifications

reported in Table 1 are referred to the HD format. The declared horizontal Field-of-View (FoV) (both for the depth and the RGB cameras) is approximately 69° and the vertical FoV is approximately 42° . The baseline between the two IR cameras is 55 mm. As can be seen from Table 1, which summarizes the technical characteristics of the device, horizontal and vertical FOVs are subject to an error of ± 3 degrees. For this reason, in Tables 2, 3 and 4 are reported the point density and framed area variations, evaluated by increasing the distance between the camera and the acquired scene. Point density variation is evaluated, respectively, for the FOV as indicated in the Intel specifications ($69.4^\circ \times 42.5^\circ$, Table 2), for the two FOV extrema of the Intel specification (i.e. obtained by removing and adding 3 degrees - Table 3), and the actual FOV of the particular camera used to carry out the performance analysis ($67.24^\circ \times 41.01^\circ$, Table 4). The FOV of the experimented camera was evaluated on a depth frame with the *get_stream_intrinsics* function implemented in the librealSense SDK.

Table 1. Technical Specifications of Intel RealSense D415.

| | |
|---|---|
| Environment | Indoor and outdoor |
| Depth Technology | Active infrared (IR) stereo |
| Image Sensor Technology | Rolling shutter: 1.4 μm x 1.4 μm pixel size |
| Depth Field of View (FOV) (Horizontal x Vertical) for HD 16:9 | $69.4 \times 42.5^\circ$ ($\pm 3^\circ$) |
| Depth Stream Output Resolution | Up to 1280 x 720 |
| Depth Stream Output Frame Rate | Up to 90 fps |
| Minimum Depth Distance (Min-Z) | 0.16 m |
| Maximum Range | ~ 10 m |
| RGB Sensor Resolution & Frame Rate | Up to 1920 x 1080 at 30 fps |
| RGB Sensor FOV (Horizontal x Vertical) | $69.4^\circ \times 42.5^\circ$ ($\pm 3^\circ$) |
| Camera Dimension (Length x Depth x Height) | 99 mm x 20 mm x 23 mm |
| Connector | USB Type-C |

Table 2. Point density and framed area variation at increasing camera-scene distance, computed using the FOV indicated in the Intel specifications ($69.4^\circ \times 42.5^\circ$).

| DISTANCE [MM] | SCALE FACTOR | X [MM] | Y [MM] | FRAMED AREA [MM ²] | POINT DENSITY [POINT / MM ²] |
|------------------|-----------------|-----------|-----------|--------------------------------------|---|
| 500 | 0,541 | 692,4 | 388,9 | 269272,400 | 3,423 |
| 600 | 0,649 | 830,9 | 466,7 | 387752,256 | 2,377 |
| 700 | 0,757 | 969,4 | 544,4 | 527773,904 | 1,746 |
| 800 | 0,866 | 1107,9 | 622,2 | 689337,344 | 1,337 |
| 900 | 0,974 | 1246,4 | 700,0 | 872442,576 | 1,056 |
| 1000 | 1,082 | 1384,9 | 777,8 | 1077089,600 | 0,856 |
| 1250 | 1,352 | 1731,1 | 972,2 | 1682952,501 | 0,548 |
| 1500 | 1,623 | 2077,3 | 1166,6 | 2423451,601 | 0,380 |

Table 3. Point density and framed area variation at increasing camera-scene distance, computed using maximum and minimum FOV variations ($69.4 \times 42.5^\circ \pm 3^\circ$).

| DISTANCE [MM] | SCALE FACTOR | X [MM] | Y [MM] | FRAMED AREA [MM ^2] | POINT DENSITY [POINT / MM ^2] | SCALE FACTOR | X [MM] | Y [MM] | FRAMED AREA [MM ^2] | POINT DENSITY [POINT / MM ^2] |
|------------------|-----------------|-----------|-----------|---------------------------|--|-----------------|-----------|-----------|---------------------------|--|
| 500 | 0,874 | 745 | 419 | 312607,4 | 0,98 | 0,748 | 638 | 359 | 229168,6 | 1,34 |
| 600 | 1,048 | 895 | 503 | 450154,7 | 0,68 | 0,898 | 766 | 431 | 330002,8 | 0,93 |
| 700 | 1,223 | 1044 | 587 | 612710,6 | 0,50 | 1,047 | 894 | 503 | 449170,5 | 0,68 |
| 800 | 1,398 | 1193 | 671 | 800275,0 | 0,38 | 1,197 | 1021 | 574 | 586671,6 | 0,52 |
| 900 | 1,573 | 1342 | 755 | 1012848,1 | 0,30 | 1,346 | 1149 | 646 | 742506,3 | 0,41 |
| 1000 | 1,747 | 1491 | 839 | 1250429,7 | 0,25 | 1,496 | 1277 | 718 | 916674,4 | 0,34 |
| 1250 | 2,184 | 1864 | 1048 | 1953796,4 | 0,16 | 1,870 | 1596 | 898 | 1432303,8 | 0,21 |
| 1500 | 2,621 | 2236 | 1258 | 2813466,8 | 0,11 | 2,244 | 1915 | 1077 | 2062517,5 | 0,15 |

Table 4. Point density and framed area variation at increasing camera-scene distance, computed using the FOV of camera under investigation (67.24°x41.01°).

| DISTANCE [MM] | SCALE FACTOR | X [MM] | Y [MM] | FRAMED AREA [MM^2] | POINT DENSITY [POINT/MM^2] |
|------------------|-----------------|-----------|-----------|-----------------------|-------------------------------|
| 500 | 0,779 | 665 | 374 | 248745,7 | 3,70 |
| 600 | 0,935 | 798 | 449 | 358193,9 | 2,57 |
| 700 | 1,091 | 931 | 524 | 487541,6 | 1,89 |
| 800 | 1,247 | 1064 | 599 | 636789,1 | 1,45 |
| 900 | 1,403 | 1197 | 673 | 805936,2 | 1,14 |
| 1000 | 1,559 | 1330 | 748 | 994982,9 | 0,93 |
| 1250 | 1,714 | 1463 | 823 | 1203929,3 | 0,77 |
| 1500 | 1,870 | 1596 | 898 | 1432775,4 | 0,64 |

3. Materials and Methods

As mentioned above, the device under examination can work in different configurations. The camera software interface supports several predefined depth presets that can be selected according to the user's usage; among the available presets, the *Default* configuration provide the best camera parameters to obtain best visual appeal, clean edges, and to reduce point cloud spraying. With the aim of providing a comprehensive characterization of the camera and to enable a comparison with the other similar devices, while keeping the analysis as general as possible, in this work the *Default* configuration is considered as the starting point from which to vary only a subset of critical parameters.

The characterization of the device is performed by considering the maximum IR cameras resolution (i.e. 1920x1080 pixels) and in a working range lower than 1 m (short range). Furthermore, tests are entirely carried out by setting the depth unit (i.e. the depth step size) to its maximum value, equal to 100 μm , to obtain the best possible depth quantization. This limit the maximum range of the camera to circa 6.5 m. Lower depth unit values are not adopted to avoid quantization effects.

The critical parameters which are varied within the experiments are the disparity shift and the laser power, the former influencing the possible working distance, the latter influencing the depth sparsity.

The D415 system evaluates the depth as a proportional inverse of the pixel disparity from right IR image to left IR image, where the pixel disparity is evaluated along the rectified epipolar lines [37]. Such a depth, named disparity shift, can be varied to modify the acquisition field.

For a given disparity shift, the maximum Z value (*MaxZ*) is given by the following formula:

$$MaxZ = focal_length(pixels) * baseline(mm) / disparity_shift, \quad (1)$$

where

$$focal_length(pixels) = \frac{1}{2} \frac{X_{res}(pixels)}{\tan(\frac{HFOV}{2})} \quad (2)$$

The minimum Z value (*MinZ*) is defined by the following equation, taking into account that the camera searches in a disparity range of 126 bit:

$$MinZ = focal_length(pixels) * baseline(mm) / (disparity_shift + 126) \quad (3)$$

By default, the disparity shift is set to a 0 value to cover the Z range from 435 mm to infinite. When the disparity shift is increased, the minimum acceptable Z value (i.e. the minimum camera-object distance) decreases, with the consequence of also reducing the maximum Z value. Figure 4 shows *MaxZ* as a function of the disparity shift, calculated considering the HFOV of the camera under investigation (67.24°). The yellow rectangle in the figure indicates the range of depth acquired based on the disparity value of the device, which is 126.

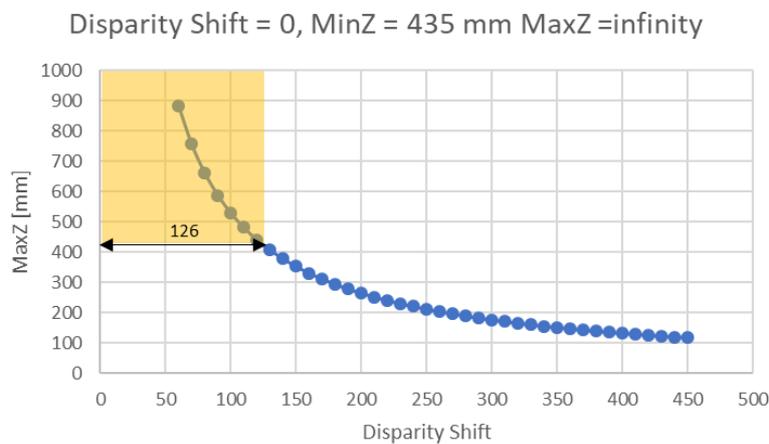


Figure 4. The graph shows the max acquirable depth as a function of the disparity shift, calculated considering the HFOV of the camera under investigation (67.24°).

As mentioned above, the second parameter used to carry out the sensor characterization consists of the laser power, i.e. the intensity of the infrared pattern projected on the scene to facilitate the search for matches between left and right image. The depth data is in fact generated with stereo vision technology that is optionally assisted by an infrared laser projector. The value of this parameter can vary in the range 0-360 mW and has a nominal value of 150 mW. If necessary, it can be increased or decreased from the nominal value for better results. For example, if a localized saturation of the laser point is noticed, the power can be reduced; to acquire distant objects, the power of the laser must be increased.

Under the premises of device setting to *Default* configuration and of using different values for disparity shift and laser power, the characterization is carried out by following the framework of Figure 5.

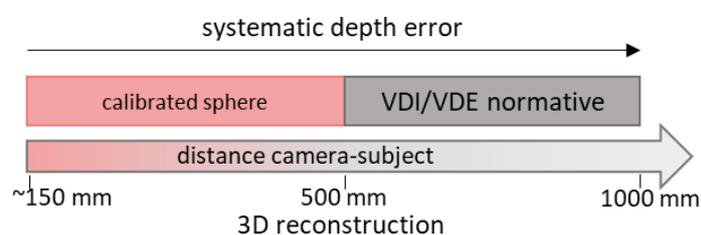


Figure 5. Proposed framework for the characterization of the device: a certified sphere is used in the very-close range (<500 mm); up to 1000 mm the VDI/VDE 2634 part 2 normative is applied; the systematic depth errors are evaluated on the entire range; the 3D reconstruction of object is tested for a distance camera-subject of 500 mm.

In detail, different tests are carried out based on the range under examination:

- From very close object acquisition (150 mm) up to 500 mm, the characterization is assessed using a calibrated sphere positioned at progressive distances from the camera with a fixed pitch of 100 mm. As already mentioned, in this range the disparity shift is required to be changed.
 - From 500 mm to 1000 mm the characterization is performed by following the guidelines of the VDI/VDE 2634 part 2 standard.
 - For the entire range a planarity test is carried out to evaluate systematic depth errors.
 - With a distance object-camera equal to 500 mm a test of the camera in allowing the three-dimensional reconstruction of objects with a multi-view acquisition is carried out. Such a reconstruction is made for two different artefacts.
- a) Characterization in the range 150-500 mm

In the short-range (i.e. from 150 to 500 mm) the VDI/VDE standard cannot be applied due to the fact that the change of the value of disparity shift does not allow to appreciate the whole working volume required from the standard. In fact, a change in the disparity shift corresponds to a variation of the working volume (which becomes smaller as soon as the target nears the sensor).

Therefore, a different method has been conceived to characterize the device in such a range, inspired by [31]. In particular, a calibrated sphere with a certified diameter of 25.4 mm was used to assess the camera's performances. The short-range was ideally divided into sub-regions with 100 mm pitch (with the exception of the range 150-200 mm) thus defining 4 operative ranges where the acquisition is performed (see Table 5). Starting from a distance between the calibrated sphere and the sensor equal to 150 mm, the sphere is acquired at increasing distances. For each operative range the disparity shift can be varied spanning from a given minimum and a maximum value to obtain the correct acquisition of the sphere, as stated in Eq. (1) and Eq. (3). This allows to define a disparity shift range for each operative range (see second column of Table 5).

The values shown in Table 5 specify the range of disparity shift that allows the acquisition in a selected depth window. The values have been calculated considering the two extremes of HFOV as indicated in the specifics (i.e. $69.4^{\circ} \pm 3^{\circ}$). Figure 6 shows an example of the calculation of the disparity shift range when considering the sub-range of depth 200-300 mm.

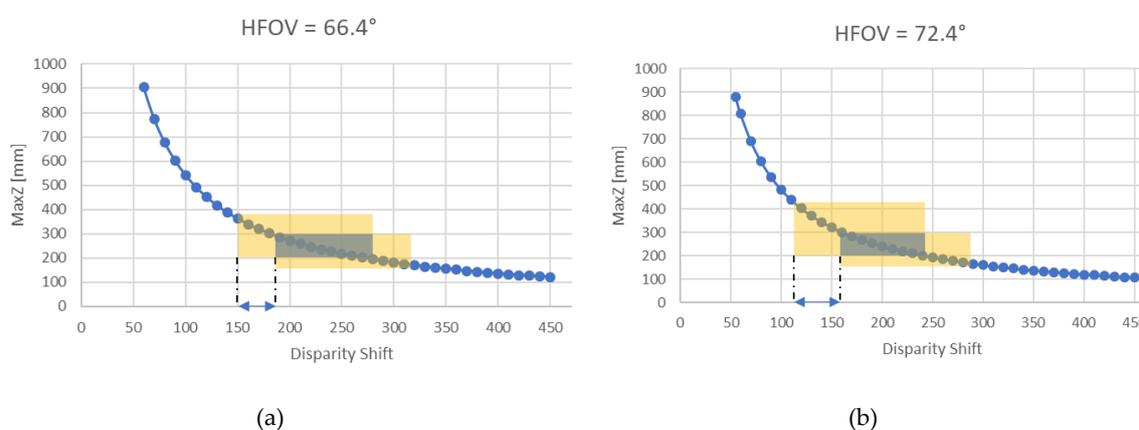


Figure 6. The graph shows the max acquirable depth as a function of the disparity shift, calculated considering $HFOV = 69.4 \pm 3^{\circ}$. The blue rectangle indicates the depth range 200-300 mm. The yellow rectangles are obtained computing the first and the last values of disparity shift that allows to acquire the whole sub-range, which are 150-180 when $HFOV = 66.4$, and 114-160 when $HFOV = 72.4$.

Figure 6a shows the case in which it is considered a HFOV equal to 66.4° , the resulting range of disparity shift is 150-180; if instead it is considered a HFOV equal to 72.4° (Figure 6b), the resulting range of disparity shift is 114-160. Consequently, if a disparity shift value between 150 and 160 is used, in both configurations the entire sub-range can be acquired.

Any value for the disparity shift falling within the disparity shift range can be selected to perform the acquisition in a given sub-region. Therefore, the preferred values are the ones listed in third column of Table 5. Using this configuration, only the laser power remains as a changing parameter to test the device performances. In addition to the nominal value, the values 0, 250 and 360 mW were tested.

Table 5. Tested value of disparity shift and laser power for each sub-range.

| Operative Range [mm] | Disparity shift range | Selected disparity shift value | Laser Power [mW] |
|-------------------------|--------------------------|-----------------------------------|---------------------|
| 150 - 200 | 234 - 240 | 235 | 0 - 150 - 250 - 360 |
| 200 - 300 | 150 - 160 | 155 | 0 - 150 - 250 - 360 |
| 300 - 400 | 49 - 125 | 50 | 0 - 150 - 250 - 360 |
| 400 - 500 | 9 - 95 | 10 | 0 - 150 - 250 - 360 |

To carry out the test, the camera was mounted on a stable tripod and the sphere was fixed on a sliding linear guide having a 4mm pitch (see Figure 7).

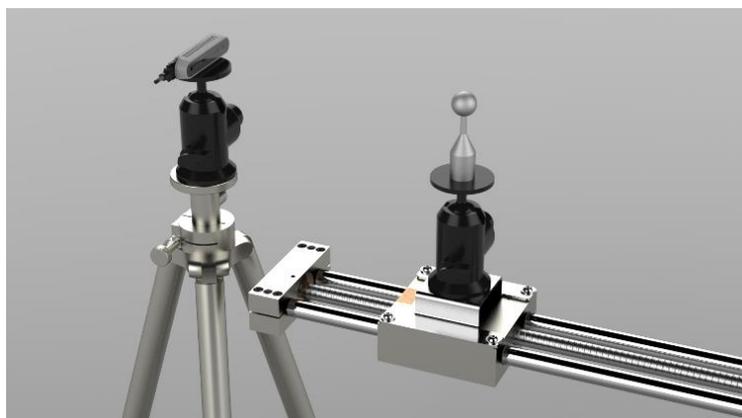


Figure 7. Acquisition setup for the very-close range evaluation test: the camera is mounted on a tripod facing the linear guide on which the sphere moves.

As said, the calibrated sphere was scanned every ten centimeters, and, in each sub-region four different laser power levels were tested (i.e. 0, 150 mW, 250 mW, 360mW). In Figure 8 two examples are reported out of the entire test set; respectively, Figure 8a shows the sphere acquired at increasing distance with maximum laser power (360 mW), and Figure 8b shows the sphere acquired at increasing distance with laser power equal to 250mW.



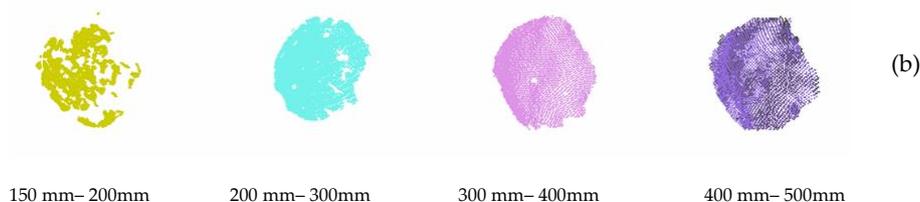


Figure 8. (a) point clouds of the sphere acquired at increasing distance with maximum laser power (360 mW); (b) point clouds of the sphere acquired at increasing distance with laser power 250mW.

Each acquisition was processed by extracting the best-fit sphere and comparing the obtained diameter (D) with the ground truth diameter ($D_{gt} = 25.4$ mm) thus defining the error E as follows:

$$E = D_{gt} - D. \quad (4)$$

Figure 9 shows the error value obtained for each sub-region with different values of the laser power parameter; such an error decreases as the laser power increases.

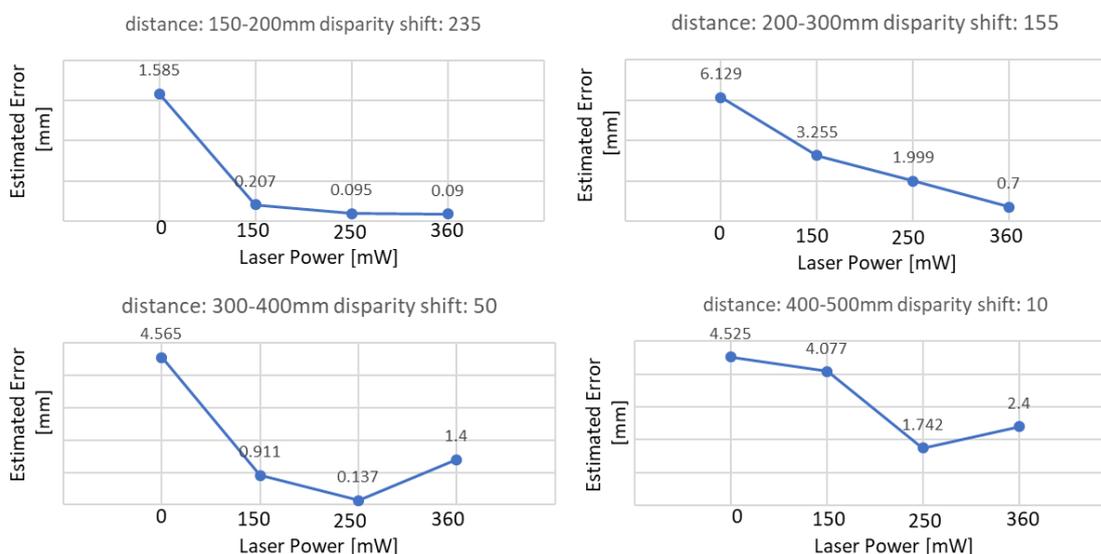


Figure 9. Errors in the estimation of the calibrated sphere for each sub-range from 150 to 500 mm, evaluated for increasing value of laser power (0 to 360 mW)

According to experimental results, in the very-close range 150-500 mm the biggest error correspond to the laser power off. When the laser power used is the default one (i.e.150 mW), the errors obtained in the reproduction of the geometry of the sphere, intended as the difference between the actual diameter and the estimated diameter, ranges from 0.2 mm and 4 mm approximately with an average error of 2.11 mm for the considered value of disparity shift. As can be seen from the graphs shown in the figure, the error decreases if the laser power value does not exceed 250 mW and it can be reduced by up to 20%; for higher values the error behavior may vary, due to the laser-speckle effect which can adversely affect the depth. The average error obtained with laser power equal to 250 mW is in fact 0.99 mm, compared to an average error of 1.15 mm obtained with laser power equal to 360 mW.

b) Characterization in the range 500-1000 mm – VDI/VDE Normative

The range 500-1000 mm has been characterized following the VDI/VDE 2634 Part 2 recommendations, in order to provide a comparison with the previous Intel camera model (the Intel RealSense SR300) and other RGB-D devices, as mentioned in Section 1. In first instance the diagonal of the working volume L_0 was determined; as mentioned above, this parameter is needed for sizing the spheres, the bar and the plane for the performance evaluation test. To this regard, the truncated pyramid representing the working volume in the range 500-1000 mm has a diagonal L_0 equal to 1291.9 mm. According to the standard, the sphere used to characterize the probing error must have diameter between 0.1 and 0.2 times L_0 . This translates into a diameter between 129.19 mm and 258.38 mm.

The distance between the centers of the two spheres of the ball-bar (to characterize the sphere-spacing error) is suggested to be greater than 0.3 times L_0 ; accordingly, in the present test such a distance is equal to 387.57 mm. Finally, the plane for the flatness error should be longer than 0.5 times L_0 , i.e. greater than 645.95 mm.

Starting from the values suggested by the standard, the artifacts chosen for carrying out the test have the following dimensions: the single sphere has a diameter of 143.15 mm, the distance between the centers of the two spheres of the ball-bar is 391.58 mm, and the plane for the flatness error is 646 mm long (Figure 10). Such values have been measured using a high precision scanner, the Romer Absolute Arm 7520 SI/SE (Hexagon Metrology S.p.A., Turin, Italy), which has an accuracy of ± 0.063 mm, and therefore allows to obtain sufficiently reliable measurements for the purpose of this work.

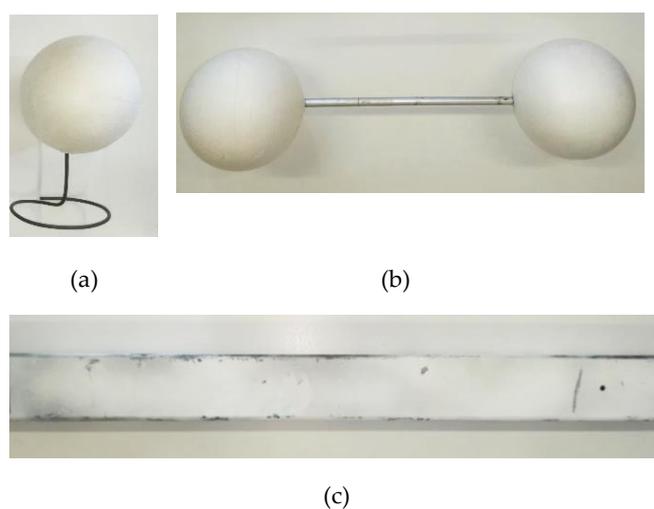


Figure 10. Artefacts proposed to measure the quality parameters inspired by the VDI/VDE normative: (a) single sphere having diameter of 143.15 mm; (b) ball-bar composed of two spheres fixed at 391.58 mm from one another; and (c) iron bar of dimensions 50x646 mm

As mentioned in Section 1, the following errors have to be measured to assess the accuracy of the imaging device.

The *form probing* error (PF) is defined as the average value of radial distances between the real measured points and a best-fit sphere evaluated according to the least-squares method, obtained for a number of acquisitions. The *size probing* error (PS) is the average value of the differences between the estimated and the “true” diameter of the sphere, again obtained using several acquisitions. To perform both the probing errors evaluation the target sphere has been positioned in 10 sequential arbitrary locations (as indicated by the standard) within the working volume (Figure 11a) thus defining 10 different values for each error. Accordingly:

$$PF_i = R_{imax} - R_{imin}, \quad (5)$$

where PF_i is the error measurement for the i^{th} acquisition, with $i = 1:10$ and R_{imax} and R_{imin} are, respectively, the maximal and minimal distance of the measured surface points of the i^{th} sphere from the center of the compensating element. Therefore, the *form probing* error is given by:

$$PF = \sum_{i=1}^{10} PF_i. \quad (6)$$

Also for the *size probing* error, it is possible to define the i^{th} measurement:

$$PS_i = D_{mi} - D_c. \quad (7)$$

Therefore, it is possible to evaluate the *size probing* error according to the following equation:

$$PS = \sum_{i=1}^{10} PS_i. \quad (8)$$

The *sphere-spacing* error SS is measured by using the balls-bar target. SS is the average value of the differences between the acquired distance and the “true” distance (lk) between the centers of the two spheres (estimated from the point cloud data using a best-fit sphere fitting) for a number of different acquisitions. The target ball-bar has been therefore positioned in 7 sequential arbitrary locations (as indicated by the standard) within the working volume (Figure 11b). For each acquisition it is possible to evaluate the j^{th} error SS_j :

$$SS_j = l_{mj} - lk. \quad (9)$$

As a consequence:

$$SS = \sum_{j=1}^7 SS_j. \quad (10)$$

The flatness (or planarity) error (F), is computed as the average of the minimum distances between two parallel planes obtained for a number of acquisitions. To measure this parameter, the plane has been sequentially uniformly-placed in 6 arbitrary positions as stated by the VDI/VDE standard within the working volume (Figure 11c) thus defining 6 F_k flatness errors.

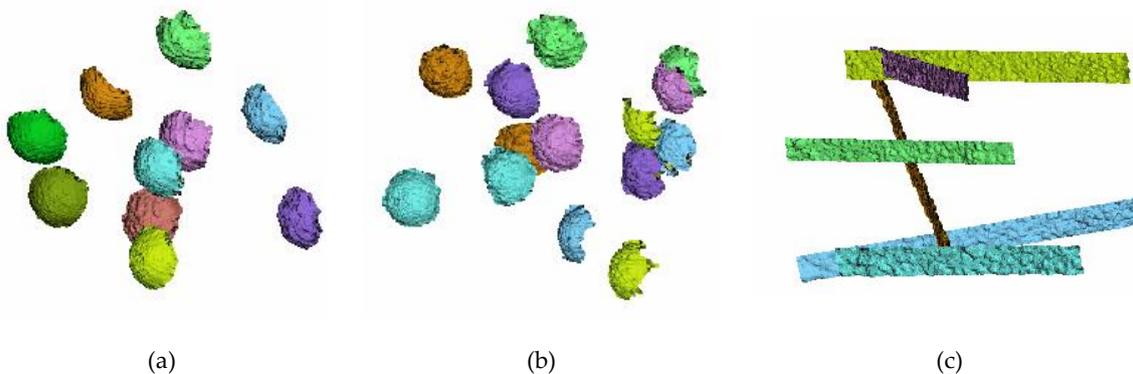


Figure 11. (a) 3D recording of the single sphere captured in 10 arbitrary positions homogeneously distributed within the working volume. (b) 3D points of the ball-bar captured in 7 positions within the working volume (the same sphere color corresponds to the same bar). (c) 3D points of the plane captured in 6 positions within the working volume.

The values obtained by calculating the errors proposed by the standard, are shown in Table 6, along with the results obtained with the Intel RealSense SR300, the Kinect v2 and the Primesense Carmine 1.09. Such results have been obtained in [33] by repeating the test setup provided by the standard for competitor cameras. As with the previous model, this experiment has been carried out on both raw and filtered data. To obtain the latter, the post-processing function of the SDK has been activated, which by default applies decimation, spatial, time and edge-preserving filtering.

Table 6. Comparison between D415, SR300, Kinect v2 and Primesense Carmine 1.09 in terms of VDI/VDE error assessment.

| Device | D415 (raw data) | D415 (filtered data) | SR 300 (filtered data) | SR300 (raw data) | Kinect v2 | Carmine 1.09 |
|------------|--------------------|-------------------------|---------------------------|---------------------|--------------|-----------------|
| P_F [mm] | 13.83 | 8.42 | 8.30 | 15.43 | 20.13 | 9.32 |
| P_S [mm] | 3.50 | 1.34 | 1.91 | 4.57 | 3.87 | 9.41 |
| SS [mm] | 4.99 | 5.03 | 6.05 | 5.11 | 19.7 | 26.08 |
| F [mm] | 15.36 | 9.45 | 6.88 | 19.33 | 12.58 | 6.71 |

Referring to probing form error P_F the new Intel device performances are comparable to the Primesense Carmine 1.09 and SR300 while Kinect v2 is characterized by a higher error. If the probing size error P_S is considered, the Intel devices with filtered data proves to be the most effective, and the Kinect v2 perform analogously to the SR300 and D415 with raw data. Higher performance of Intel devices, both D415 and SR300 is even more evident when dealing with sphere-spacing error SS; in fact, the Δ_{SS} is considerably higher for tested competitor cameras. The performance in terms of flatness error F is almost the same for the SR300 and the D415 with filtered data and Primesense Carmine 1.09. Interestingly, the behavior of Kinect v2 in all tests is quite satisfying considering that this camera system is specifically designed to work in medium to long range. It is interesting to note that the sphere spacing error measured for SR300 and D415 using raw data is almost the same assessed using optimized settings. This may be caused by the moderate smoothing effect obtained using the optimized setting does not have particular effect on the position of spheres centrum when compared with the same position evaluated using raw data. When comparing the two RealSense models, it is noted that the results obtained are comparable: if the raw data are considered the new camera shows slightly better performances on the whole test-set; on the other hand, the results obtained with filtered data are similar.

As reported in [33] the discussion can be extended to other results from scientific literature: [38] report the results computed for the probing errors (P_F and P_S) with several cameras, among them the Occipital Structure Sensor device, which is one of the most relevant competitor of Intel RealSense devices. This sensor scored an average value of ~ 10 mm for P_F and ~ 2.2 mm for P_S , while the values obtained with the latest intel model are 13.83 mm for raw data and 8.42 mm for filtered data, for P_F , and 3.5 mm and 1.91 mm for P_S .

Also regarding the flatness error F, the performances of the D415 can be compared to the Occipital Structure Sensor device [38] with the result of reducing the error up to 50% when using D415 with filtered data.

The sphere spacing error SS can instead be compared to the results in [39] obtained with the Asus Xtion Pro camera, which results span from -8 to 2 mm against ~ 5 mm obtained with the D415 model.

c) Systematic depth errors in the entire range 150-1000 mm

As also reported in [33], one of the most common systematic errors for RGB-D devices is the so-called "inhomogeneous distance", i.e. the set of possible errors that may result from acquisitions at different distances from the sensor. To study these errors, the camera is positioned perpendicularly to a plane that is acquired at incremental distances (see Figure 12a) in the range 200 - 1000 mm, to

maintain a regular pitch of 100mm. In the test the perpendicularity has been ensured by inserting a reference mark in the scene in correspondence with the center of the image, verifying that the point is maintained at the center as the camera moves away from the plane (Figure 12b). The plane-camera distance is assured by the pitch of the linear guide on which it has been mounted (Figure 12a) that results to be equal to 4mm, therefore known a priori.

The aim is to assess two types of errors: systematic *non-planarity* errors and *offset* errors on the depth. The first one was evaluated by referring to a ground truth plane (with certified flatness of 50 μm), as in [33], to verify the flatness of the obtained scan. The latter was analyzed, studying the distance in Z between the obtained scan and the ground truth plane.

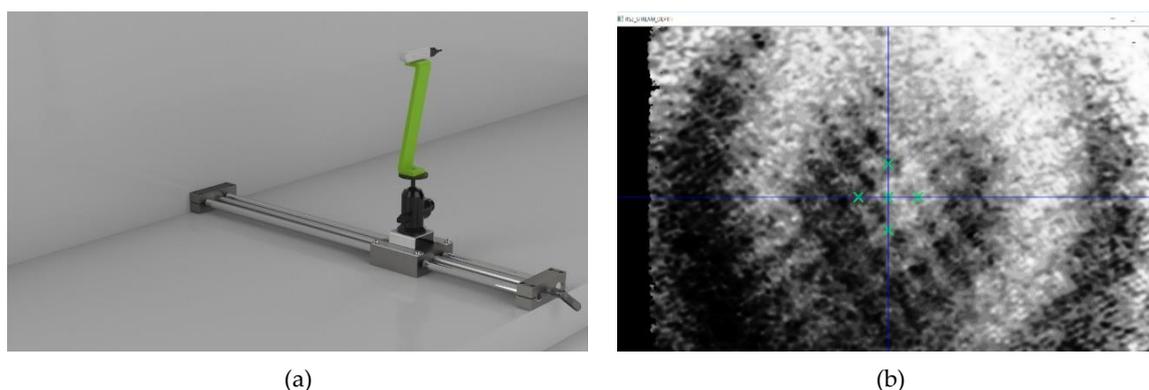


Figure 12. (a) Acquisition setup for systematic depth errors evaluation test: the camera is mounted on a linear guide perpendicular to a planar surface. (b) The calibration of the optical axis of the camera has been assessed by framing the planar surface and verifying that the IR image center (green cross) corresponds to the depth map center for closer and further distances.

Figure 13 shows the results obtained for the two tests. In detail, Figure 13a shows the error map between the scanned data and the best fitting planes built on such data, to highlight the non-planarity error that is introduced by the camera. The average deviation of the point clouds with respect to the ground truth planes ranges from 0.004 mm with a standard deviation of 0.345 mm at a distance of 200 mm from the plane, up to 0.237 mm with a standard deviation of 4.74 mm at a distance of one meter. The maximum recorded error value spans from 1.6 mm to 26.5 mm, respectively at 200 mm and at 1000 mm distance. Figure 13b shows, instead, that nearly no error is present with regard to the depth offset, i.e. the point clouds are almost correctly positioned at a distance of 100 mm from each other. In fact, the best fitting planes evaluated for each point cloud using the Geomagic Design X® software package (3D Systems, Rock Hill, South Carolina, USA), show a distance from the ground truth with a minimum value of 0.06 mm and a maximum value of 2.54 mm for the planes acquired at distances of 1000 mm and 900 mm respectively.

The analysis of systematic depth errors led to interesting results, especially when compared against the ones obtained for the SR300. In first instance, compared to the previous model, both the non-planarity of the single acquisitions and the depth offset have improved considerably. Specifically, as shown in Figure 13b, there is nearly no offset between the acquired data and the best fitting planes; moreover, the clouds do not have the “twist effect” that was noticed of the previous model. Noise tends to be distributed locally leading to higher peaks than those reported by the previous sensor; it can also be noted a non-linearity probably introduced by the lenses [31].

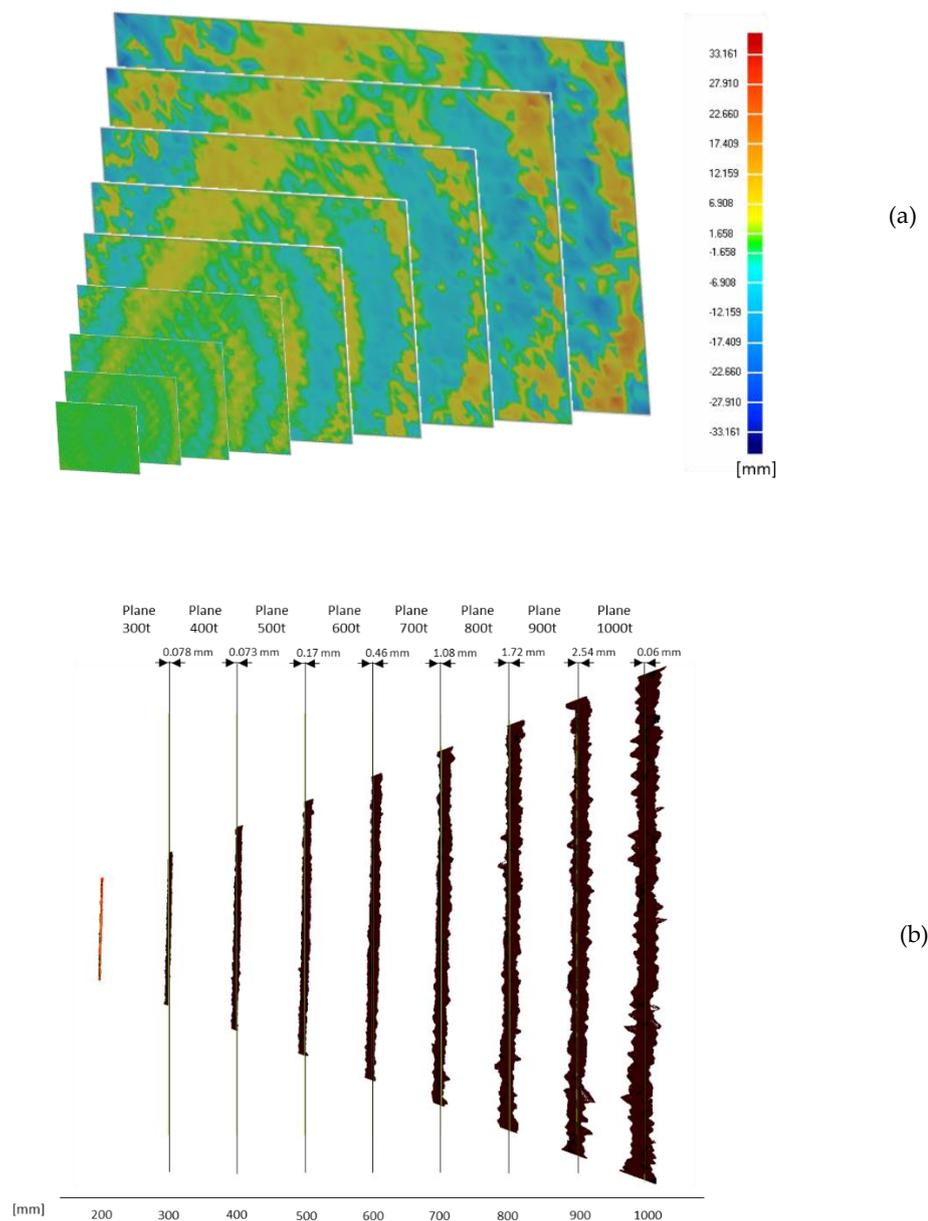


Figure 13. (a) Error map between the target plane and the 9 point clouds acquired using the D415 device with increasing distance from the plane itself. (b) Evaluation of the depth-offset error for the 9 acquired planes.

d) 3D object reconstruction

The last test carried out concerns the three-dimensional reconstruction of objects. Inspired by part 3 of the VDI/VDE 2634 standard [40], which provides guidelines for the positioning of the camera around the artifact, the object under examination was rotated with respect of a stable camera.

It was decided to evaluate the accuracy of the reconstruction with the same two objects used to characterize the SR300 in [33]: a freeform smooth object (~ 200 mm height and ~ 70 mm wide) and a 3D Tangram object (bounding box with dimensions 150 mm height and 150 mm wide). Again, to provide an effective and easily replicable evaluation, the camera has been used in the default configuration, being the only modified parameter the depth unit, set to the best achievable resolution. To ensure the best performances the camera-object distance has been set to the smaller achievable value of approximately 500 mm. To obtain two reference models for assessing the device

performances, the Romer Absolute Arm 7520 SI/SE has been used (see Figure 14). The meshes, obtained using Geomagic® Studio software package starting from the point cloud acquired with the D415 camera, and the reference 3D model have been globally registered using the ICP algorithm [41]. To increase the quality of the mesh reconstruction, 5% borders have been removed prior to perform the alignment.



Figure 14. (a) Test objects for multiple-view reconstruction: a freeform statue and a 3D tangram; (b) reference 3D models for multiple-view reconstruction, obtained starting from the Romer Absolute Arm acquisition (accuracy of ± 0.063 mm).

Figure 15 shows the Euclidean distances between the acquired artifacts and the ground truth. The comparison between the ground truth and the acquired data is limited to the target portion obtained with the multiple-view acquisition; grey areas in Figures 15 are not, therefore, considered. As far as the tangram (Figure 15a) is concerned, an average deviation of 0.110 mm is obtained with a standard deviation of 0.902 mm; for what concerns the statue (Figure 15b), the average deviation obtained is -0.029 mm with a standard deviation of 1.856 mm.

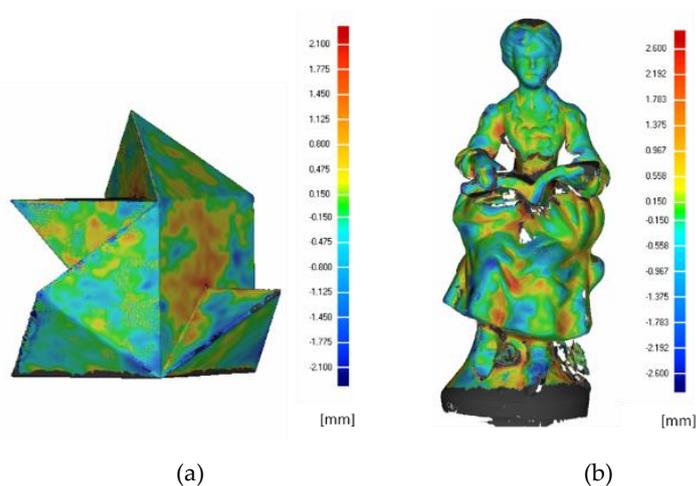


Figure 15. (a) 3D comparison between the tangram ground truth and the D415 acquired data. Errors are in mm. (b) 3D comparison between the statue ground truth and the D415 acquired data. Errors are in mm.

The multi-views acquisition test for the full 3D reconstruction proves (see Figure 15) that both geometries are fairly measured, particularly if it is considered that for other RGB-D devices (e.g. the Microsoft Kinect) it has been evaluated that 3D reconstruction error can span in an interval of 27 mm [34]. It can also be noted that, there is no substantial difference with regard to the error assessed in the two reconstructions of the prismatic shape (tangram 3D) and the free form shape (statue), when the latest Intel model is used. Furthermore, the results shown here can be compared with the results

obtained by scanning the same objects with the SR300 (Table 7). It can be noted that the results are fully comparable, the main difference is found in the reconstruction of the statue which has, as said, a free form shape: the new model perform an average error of -0.029 mm against -0.2 mm obtained with the previous model.

Table 7. Average deviation (avg) and standard deviation (std) from the ground truth data obtained for the 3D reconstruction of the two artifacts for D415 and SR300 cameras

| | Reconstruction deviation from ground truth | |
|-------|--|--------------------------------|
| | Tangram | Statue |
| D415 | avg: 0.11 mm, std:0.90 mm | avg: -0.029 mm, std:1.85 mm |
| SR300 | avg: 0.10 mm, std: 0.4 mm | avg: -0.2 mm, std: 1.01 mm |

4. Discussion and Conclusion

Given the growing popularity of RGB-D devices mainly due to their versatility, and the interest that Intel RealSense devices have obtained in recent years for their accuracy, compactness and ease of use, in this work a metrological evaluation of the latest model presented by Intel, the RealSense D415, was carried out. After investigating the operating principles of this device, a panel of tests was defined, based on the distance between the camera and the examined scene. The main objective was, other than characterizing the new model, to compare this device with the previous model presented by Intel, the RealSense S300 and where possible with other competitor RGB-D cameras.

In this work the devices were compared through 4 types of tests: the error measured using a calibrated sphere in the very close range, the errors measured with the VDI/VDE 2634 part 2 normative, the systematic depth errors extracted with the acquisition of a planar surface at increasing distances and the 3D reconstruction of object.

Tests have shown that the new D415 model is fully comparable to its predecessor model in terms of errors assessed through VDI/VDE normative. The device is also in line with the results obtained with other devices in the scientific literature and exceeds their performance when considering the filtered data.

As for the 3D reconstruction, a comparison can be made with the previous Intel model, since they both work in the close-range with the result of a comparable deviation error.

The most interesting result is obtained from the estimation of the systematic depth errors: the D415 in fact reports better results both in terms of average flatness error, and of displacement in Z from the real plane with a maximum displacement from the ground truth planes of 2.52 mm.

In addition, the device is sufficiently accurate when acquiring very close distances (e.g. from 150 mm to 500 mm) as the reconstruction error in this range can be 0.99 mm on average when using a laser power equal to 250 mW.

As demonstrated by the experimental results of the characterization, it is possible to state that despite the device is born for addressing applications such as tracking, gaming or gesture recognition, it could also be satisfactorily employed as a 3D scanner i.e. could be referred as a low-cost device for a number of 3D scanning applications dealing with, among others, health, fashion, fitness and Cultural Heritage.

Funding: This research received no external funding

Conflicts of Interest: The authors declare no conflict of interest.

References

- [1] Sturm J, Engelhard N, Endres F, Burgard W, Cremers D. A benchmark for the evaluation of RGB-D SLAM systems. 2012 IEEE/RSJ Int. Conf. Intell. Robot. Syst., IEEE; 2012, p. 573–80.

- doi:10.1109/IROS.2012.6385773.
- [2] Makris S, Tsarouchi P, Surdilovic D, Krüger J. Intuitive dual arm robot programming for assembly operations. *CIRP Ann* 2014;63:13–6. doi:10.1016/j.cirp.2014.03.017.
 - [3] Prabhu VA, Muhandri N, Song B, Tiwari A. Decision support system enabled by depth imaging sensor data for intelligent automation of moving assemblies. *Proc Inst Mech Eng Part B J Eng Manuf* 2018;232:51–66. doi:10.1177/0954405416645984.
 - [4] Bakhtiyari K, Beckmann N, Ziegler J. Contactless heart rate variability measurement by IR and 3D depth sensors with respiratory sinus arrhythmia. *Procedia Comput Sci* 2017;109:498–505. doi:10.1016/j.PROCS.2017.05.319.
 - [5] Carfagni M, Furferi R, Governi L, Servi M, Uccheddu F, Volpe Y, et al. Fast and Low Cost Acquisition and Reconstruction System for Human Hand-wrist-arm Anatomy. *Procedia Manuf* 2017;11. doi:10.1016/j.promfg.2017.07.306.
 - [6] Buonamici F, Furferi R, Governi L, Lazzeri S, McGreevy KS, Servi M, et al. A Practical Methodology for Computer Aided Design of Custom 3D Printable Casts for Wrist Fractures. *Vis Comput* 2018.
 - [7] Primesense is changing the way we interact with digital devices PrimeSense™ 3D Sensors ®. n.d.
 - [8] Xtion PRO | 3D Sensor | ASUS Global n.d. https://www.asus.com/3D-Sensor/Xtion_PRO/ (accessed November 27, 2018).
 - [9] Astra Pro – Orbbec n.d. <https://orbbec3d.com/product-astra-pro/> (accessed November 27, 2018).
 - [10] Structure Sensor Press Info. Structure n.d. <https://structure.io/press#press-photos> (accessed May 4, 2018).
 - [11] Gao X, Zhang T. Robust RGB-D simultaneous localization and mapping using planar point features. *Rob Auton Syst* 2015;72:1–14. doi:10.1016/j.robot.2015.03.007.
 - [12] Kinect - Windows app development n.d. <https://developer.microsoft.com/en-us/windows/kinect> (accessed November 27, 2018).
 - [13] Lenovo Phab 2 Pro | Augmented Reality Smartphones | Lenovo US n.d. <https://www.lenovo.com/us/en/smart-devices/-lenovo-smartphones/phab-series/Lenovo-Phab-2-Pro/p/WMD00000220> (accessed December 11, 2018).
 - [14] iPhone - Apple n.d. <https://www.apple.com/lae/iphone/> (accessed December 11, 2018).
 - [15] Zhang S. High-speed 3D shape measurement with structured light methods: A review. *Opt Lasers Eng* 2018;106:119–31. doi:10.1016/j.optlaseng.2018.02.017.
 - [16] Keselman L, Woodfill JI, Grunnet-Jepsen A, Bhowmik A. Intel R RealSense™ Stereoscopic Depth Cameras. n.d.
 - [17] Biometric Facial Recognition – Windows Hello - Microsoft n.d. <https://www.microsoft.com/en-us/windows/windows-hello> (accessed December 11, 2018).
 - [18] Arachchi SPK, Hakim NL, Hsu H-H, Klimenko SV, Shih TK. Real-Time Static and Dynamic Gesture Recognition Using Mixed Space Features for 3D Virtual World's Interactions. 2018 32nd Int. Conf. Adv. Inf. Netw. Appl. Work., IEEE; 2018, p. 627–32. doi:10.1109/WAINA.2018.00157.
 - [19] Jones L. Your Body of Water: A Display that Visualizes Aesthetic Heart Rate Data from a 3D Camera 2018. doi:10.1145/3173225.3173284.
 - [20] Liao B, Li J, Ju Z, Ouyang G. Hand Gesture Recognition with Generalized Hough Transform and DC-CNN Using Realsense. 2018 Eighth Int. Conf. Inf. Sci. Technol., IEEE; 2018, p. 84–90. doi:10.1109/ICIST.2018.8426125.
 - [21] Chiang T, Fan C-P. 3D Depth Information Based 2D Low-Complexity Hand Posture and Gesture Recognition Design for Human Computer Interactions. 2018 3rd Int. Conf. Comput. Commun. Syst.,

- IEEE; 2018, p. 233–8. doi:10.1109/CCOMS.2018.8463327.
- [22] Karambakhsh A, Kamel A, Sheng B, Li P, Yang P, Feng DD. Deep gesture interaction for augmented anatomy learning. *Int J Inf Manage* 2018. doi:10.1016/J.IJINFOMGT.2018.03.004.
- [23] Yoon Ahn J, Wan Kim D, Hyeon Lee Y, Kim W, Kuk Hong J, Shim Y, et al. MOYA: Interactive AI toy for children to develop their language skills 2018. doi:10.1145/3174910.3174957.
- [24] Bock R. Low-cost 3D security camera. In: Dudzik MC, Ricklin JC, editors. *Auton. Syst. Sensors, Veh. Secur. Internet Everything*, vol. 10643, SPIE; 2018, p. 15. doi:10.1117/12.2305455.
- [25] Fang Q, Kyrarini M, Ristić-Durrant D, Gräser A. RGB-D Camera based 3D Human Mouth Detection and Tracking Towards Robotic Feeding Assistance 2018. doi:10.1145/3197768.3201576.
- [26] Lasso A, Ungi T, Fichtinger G, Asselin M. Towards webcam-based tracking for interventional navigation. In: Webster RJ, Fei B, editors. *Med. Imaging 2018 Image-Guided Proced. Robot. Interv. Model.*, vol. 10576, SPIE; 2018, p. 68. doi:10.1117/12.2293904.
- [27] Siena FL, Byrom B, Watts P, Breedon P. Utilising the Intel RealSense Camera for Measuring Health Outcomes in Clinical Research. *J Med Syst* 2018;42:53. doi:10.1007/s10916-018-0905-x.
- [28] Aoki H, Suzuki A, Shiga T. Study on Non-contact Heart Beat Measurement Method by Using Depth Sensor, Springer, Singapore; 2019, p. 341–5. doi:10.1007/978-981-10-9035-6_62.
- [29] Yu X, Yang G, Jones S, Sanjie J. AR Marker Aided Obstacle Localization System for Assisting Visually Impaired. 2018 IEEE Int. Conf. Electro/Information Technol., IEEE; 2018, p. 0271–6. doi:10.1109/EIT.2018.8500166.
- [30] Intel® RealSense™ Cross Platform API (librealsense): Main Page n.d. <https://software.intel.com/sites/products/realsense/camera/> (accessed November 27, 2018).
- [31] Giancola S, Valenti M, Sala R. A Survey on 3D Cameras: Metrological Comparison of Time-of-Flight, Structured-Light and Active Stereoscopy Technologies. n.d.
- [32] VDI/VDE 2634 Blatt 2 n.d. https://www.vdi.eu/guidelines/vdivde_2634_blat_2-optische_3_d_messsysteme_bildgebende_systeme_mit_flaechenhafter_antastung/ (accessed November 28, 2018).
- [33] Carfagni M, Furferi R, Governi L, Servi M, Uccheddu F, Volpe Y. On the Performance of the Intel SR300 Depth Camera: Metrological and Critical Characterization. *IEEE Sens J* 2017;17:4508–19. doi:10.1109/JSEN.2017.2703829.
- [34] Lachat E, Macher H, Landes T, Grussenmeyer P. Assessment and Calibration of a RGB-D Camera (Kinect v2 Sensor) Towards a Potential Use for Close-Range 3D Modeling. *Remote Sens* 2015;7:13070–97. doi:10.3390/rs71013070.
- [35] Guidi G, Gonizzi S, Micoli L. 3D CAPTURING PERFORMANCES OF LOW-COST RANGE SENSORS FOR MASS-MARKET APPLICATIONS. *ISPRS - Int Arch Photogramm Remote Sens Spat Inf Sci* 2016;XLI-B5:33–40. doi:10.5194/isprsarchives-XLI-B5-33-2016.
- [36] Chiabrando F, Chiabrando R, Piatti D, Rinaudo F, Chiabrando F, Chiabrando R, et al. Sensors for 3D Imaging: Metric Evaluation and Calibration of a CCD/CMOS Time-of-Flight Camera. *Sensors* 2009;9:10080–96. doi:10.3390/s91210080.
- [37] Morgan GLK, Liu JG, Yan H. Precise Subpixel Disparity Measurement From Very Narrow Baseline Stereo. *IEEE Trans Geosci Remote Sens* 2010;48:3424–33. doi:10.1109/TGRS.2010.2046672.
- [38] Kersten TP, Lindstaedt M, Starosta D. COMPARATIVE GEOMETRICAL ACCURACY INVESTIGATIONS OF HAND-HELD 3D SCANNING SYSTEMS – AN UPDATE. *ISPRS - Int Arch Photogramm Remote Sens Spat Inf Sci* 2018;XLII-2:487–94. doi:10.5194/isprs-archives-XLII-2-487-2018.

- [39] Boehm J. Accuracy Investigation for Structured-light Based Consumer 3D Sensors. *Photogramm - Fernerkundung - Geoinf* 2014;2014:117–27. doi:10.1127/1432-8364/2014/0214.
- [40] VDI/VDE 2634 Blatt 3 n.d. https://www.vdi.eu/guidelines/vdivde_2634_blat_3-optische_3_d_messsysteme_bildgebende_systeme_mit_flaechenhafter_antastung/ (accessed November 28, 2018).
- [41] Zhang Z. Iterative Point Matching for Registration of Free-Form Curves and Surfaces. vol. 13. 1994.