

Review

# Perspectives of RealSense and ZED Depth Sensors for Robotic Vision Applications

Vladimir Tadic <sup>1,\*</sup>, Attila Toth <sup>2</sup>, Zoltan Vizvari <sup>3</sup>, Mihaly Klincsik <sup>4</sup>, Zoltan Sari <sup>4</sup>, Peter Sarcevic <sup>5</sup>, Jozsef Sarosi <sup>5</sup> and Istvan Biro <sup>5</sup>

<sup>1</sup> Institute of Information Technology, University of Dunaujvaros, Tancsics Mihaly u. 1/A Pf.: 152, H-2401 Dunaujvaros, Hungary

<sup>2</sup> Institute of Physiology, Medical School, University of Pecs, Szigeti Str. 12, H-7624 Pecs, Hungary; attila.toth@aok.pte.hu

<sup>3</sup> Department of Environmental Engineering, Faculty of Engineering and Information Technology, University of Pecs, Boszorkany Str. 2, H-7624 Pecs, Hungary; vizvari.zoltan@mik.pte.hu

<sup>4</sup> Department of Technical Informatics, Faculty of Engineering and Information Technology, University of Pecs, Boszorkany Str. 2, H-7624 Pecs, Hungary; klincsik@mik.pte.hu (M.K.); sari.zoltan@mik.pte.hu (Z.S.)

<sup>5</sup> Faculty of Engineering, University of Szeged, Mars ter 7, H-6724 Szeged, Hungary; sarcevic@mk.u-szeged.hu (P.S.); sarosi@mk.u-szeged.hu (J.S.); biro-i@mk.u-szeged.hu (I.B.)

\* Correspondence: laslo.tadic@gmail.com

**Abstract:** This review paper presents an overview of depth cameras. Our goal is to describe the features and capabilities of the introduced depth sensors in order to determine their possibilities in robotic applications, focusing on objects that might appear in applications with high accuracy requirements. A series of experiments was conducted, and various depth measuring conditions were examined in order to compare the measurement results of all the depth cameras. Based on the results, all the examined depth sensors were appropriate for applications where obstacle avoidance and robot spatial orientation were required in coexistence with image vision algorithms. In robotic vision applications where high accuracy and precision were obligatory, the ZED depth sensors achieved better measurement results.

**Keywords:** depth sensors; depth map; RealSense; ZED; ZED 2i; robotic applications



**Citation:** 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* **2022**, *10*, 183. <https://doi.org/10.3390/machines10030183>

Academic Editor:

Antonios Gasteratos

Received: 28 January 2022

Accepted: 1 March 2022

Published: 3 March 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Not long ago, robotic vision applications based on two-dimensional (2D) image processing were largely limited because of an absence of information on the depth in the direction of the Z coordinate. In contrast to 2D computer vision, three-dimensional (3D) image vision enables computers and other devices to discern accurately and precisely various distances and shapes, as well as to control certain robots in the real, 3D world [1–20]. For this purpose, 3D optical systems have been successfully implemented in multifarious scientific fields such as robotics, car manufacturing, automatics, mechanics, biomedicine, surveillance systems, industry, etc. [21–40].

Up to now, 3D depth cameras have been very expensive, and their utilization has been complex and burdensome from a hardware standpoint. Nowadays, thanks to technical progress, the price of 3D depth sensors that can measure image depth is considerably affordable and their use much simpler [41–59].

The first version of the Kinect sensor was launched in 2009 by the company PrimeSense. The Kinect contained an infrared (IR) lens, and its sensors detected IR points projected onto a scene, whose depth information was appraised along the Z coordinate. In the first version, the resolution of the depth sensor was only  $320 \times 240$  pixels with 2048 levels of depth. Later, other manufacturers constructed and manufactured depth sensors similar to the Kinect sensor that were typically meant for home use, primarily in interactive video games, at reasonable prices [1].

In 2013, Apple vested the PrimeSense company after it developed and manufactured the Kinect camera, and its newer sensors became much more accurate and precise in determining the depth information in an image. Recently, some depth cameras have also become an integral part of smartphones [1].

In the meantime, Intel also began deploying its own series of depth cameras, the RealSense depth sensors. The first version of these sensors was developed in cooperation with Microsoft Company for the Windows operating system as part of the face recognition login system for Windows 10 [1–7].

Although RealSense sensors only recently appeared on the market, they rapidly found their application in numerous fields such as robotics, medicine, face recognition, interactive children's toys, recognition and tracking human movements, etc. [7–15].

Finally, numerous depth sensor manufacturers have recently begun producing commercial and industrial depth sensors, and a major company in this field is Stereolabs with its ZED depth sensor series. ZED depth sensors are considered to be one of the best commercial depth sensors on the market [16–20].

This comprehensive review paper is a response to an invitation from the guest editors to contribute a perspective overview-based paper on depth sensors for a Special Issue related to modeling, sensor fusion, and control techniques in applied robotics.

The goal of this research is to investigate the capabilities of available depth sensors in order to determine their abilities for the development of low-cost robotic applications that require some kind of robotic vision algorithm. Namely, the hardware required for robotic vision can be very expensive if industrial Red-Green-Blue (RGB) cameras, industrial depth sensors and industrial 3D scanners are used. The high cost of an imaging equipment can lead to the fact that many investors may abandon the robotization of certain applications because of the high cost of only one component of the system. Therefore, the aim of this paper is to examine how available low-cost commercial depth sensors could affect the development of certain low-budget robotic applications.

Four depth sensors were initially procured, whereas the newest ZED 2i depth camera was procured recently, and these four depth sensors served as the basis for this research and paper. These depth cameras were purchased for research purposes to determine which depth camera could be used in research related to industrial and medical robotic applications in the near future. We hope that this paper will guide and assist readers and researchers who would like to take advantage of depth sensors in low-cost robotic vision applications in the future.

In this paper, the D415 and D435 of the D400 series of Intel RealSense sensors and Stereolabs' ZED and ZED 2i depth sensors will be described alongside depth sensing experiments and comparison of examples. The available depth cameras will be tested in various situations in order to determine their capabilities.

The paper is organized as follows: Section 1 is the introduction with a related brief historical overview of the depth camera; Section 2 summarizes the RealSense depth sensors, and Section 3 introduces the ZED cameras. Section 4 describes the experiments and results with a comparison of the cameras' capabilities, while Section 5 provides the conclusion, followed by future works in Section 6.

## 2. RealSense Depth Sensors

This section will present the technology and some of the more important properties of the Intel RealSense depth sensors along with their working principle.

RealSense technology comprises a microprocessor for image processing, a module for creating depth images, an IR emitter, a segment for tracking movements, and depth sensors. These depth sensors are built on deep scanning technology, which allows devices to see shapes and objects in the same manner as humans. The complete hardware is also supported by appropriate open-source Software Development Kit (SDK) software called *librealsense* [7]. This software platform provides simple software support for all RealSense cameras. The software platform supports C/C++, ROS (Robot Operating System), Python,

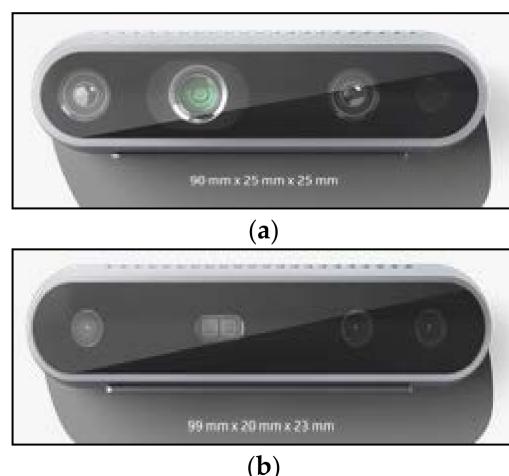
MATLAB, etc., systems and programming languages for the development of appropriate and various applications. Intel also provides two applications for the setup and use of the cameras [8].

Intel introduced the D400 series RealSense cameras in 2018 with the D415 and D435 models. By presenting these sensors, Intel became an important manufacturer on the market in terms of a balance of quality and price.

These RealSense sensors principally differ in the field of view (FOV) measured in angles and the type of shutter that tunes the exposure.

The D435 camera has a wider FOV ( $H \times V \times D$ —Horizontal  $\times$  Vertical  $\times$  Diagonal):  $91^\circ \times 65^\circ \times 100^\circ$  for RGB camera, which minimizes blind, black spots in the depth map, after which the acquired depth image is pleasing to the eye. The FOV of the corresponding depth sensor is ( $H \times V \times D$ ):  $85^\circ \times 58^\circ \times 90^\circ$ . As a result, this depth sensor is suitable for applications where no great accuracy and precision are required, but where a global visual experience is more significant. Accuracy is the percentage of error with respect to the measured depth, while precision is the capability of the sensors to replicate the same measure in the same conditions [16]. These are important parameters of depth sensors, and they can be easily observed by observing the measured depth maps of a scene, or object [7]. Usually, visual assessment is used in practice for the characterization of resulted depth images [20,21]. Therefore, this sensor is often used in automotive applications and in drones. Furthermore, this sensor has a global shutter that ensures a better performance in situations where lighting is unsatisfactory, while capturing fast movements in a scene, and reduces the effect of blurring in images [7]. The D435 depth camera also yields better depth measurement results when the targeted objects are a few meters away from the camera itself.

The D415 depth camera has a narrower FOV ( $H \times V \times D$ ):  $69^\circ \times 42^\circ \times 77^\circ$  for RGB camera, and this property results in a higher density of pixels, thus increasing the resolution of the depth map. Here, the FOV of the corresponding depth sensor is ( $H \times V \times D$ ):  $65^\circ \times 40^\circ \times 70^\circ$ . Hence, if accuracy and precision are the main requirements in an application, e.g., avoiding obstacles in robotics application and in object detection, the D415 depth sensor gives much better results. The RealSense D415 sensor has a rolling shutter. This attribute improves the performance of this depth sensor when there are no unexpected fast movements during image capturing, but the image is static [7,11]. It should be noted that the D415 depth camera yields better depth measurement results when the targeted object is close to the sensor, i.e., 1 m or less. Both the D435 and D415 cameras yield the best quality depth maps at about 60–70 cm from the scene, according to the literature [7,11]. Figure 1 shows the RealSense depth cameras [7].



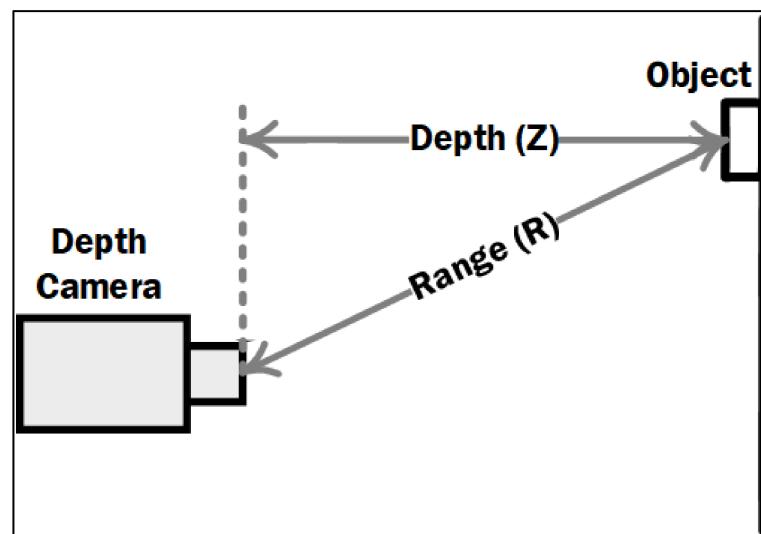
**Figure 1.** (a) Intel RealSense D435 and (b) Intel RealSense D415 cameras [7].

Thanks to the RealSense SDK software kit, the user interface of the built-in applications provides the highest control levels, which was unthinkable until recently [11].

The RealSense depth sensors have three camera lenses: an IR camera, a RGB camera, and an IR laser projector. Hence, these depth sensors are called active devices since they contain a ranging IR laser projector in order to improve the depth measurement. All three lenses in conjunction enable it to assess the depth information by detecting the IR light that is reflected from the object in front of the lenses. The resulting visual information, combined with the SDK software, generates a depth estimate, i.e., produces a depth map. After further post-processing, a depth image yielded in this way can be used, for example, for tracking movements or detecting objects, by creating a user interface that gives the impression of touch, which reacts to movements of the head, the leg, the hand, or any other body part. Naturally, since the RealSense depth sensor also possess an RGB lens and an IR lens, it is therefore possible to capture images in color and in conditions of poor lighting [7,11].

RealSense sensors use stereovision to calculate depth [7]. The realization of stereovision consists of a right-side and a left-side sensor and of an IR projector. The IR projector projects invisible IR rays that improve the accuracy and precision of the depth data in scenes with poor textures. The right-side and left-side sensors capture the scene and send information about the real image to the microprocessor. Based on the received data, the image processor determines via stereometry calculation the depth values for each pixel of the recorded image, thus correlating the values obtained with the right-side camera to the image derived with the left-side camera. The depth data of each pixel processed in this manner result in the depth image. By linking up successive depth images, a depth video stream is generated.

As seen in Figure 2, the value of the depth pixel representing the depth ( $Z$ ) of an obstacle/object is determined in relation to a parallel plane of the depth sensor doing the capturing and not in relation to the actual range ( $R$ ) of the obstacle/object from the depth sensor.



**Figure 2.** Diagram of the depth measuring method in relation to depth and range [7].

A crucial role in the operation of the depth sensor is also played by the RealSense D4 digital signal processor (DSP) for image processing [7]. This DSP processor can process 36 million depth values in a second. Thanks to this high performance, these depth sensors are built into a multitude of electronic devices that necessitate high-speed data processing [7,11].

According to Intel [7], the main features of RealSense depth devices are summarized and compared in Table 1.

**Table 1.** Features of D415 and D435 depth cameras [7].

Features	D415	D435
Depth resolution	16 bit	16 bit
Max. RGB resolution	1920 × 1080 pixels	1920 × 1080 pixels
Range	0.2 m–10 m	0.2 m–10 m
Diagonal field of view	70°	90°
Shutter	rolling shutter	global shutter

### 3. ZED Depth Sensors

This section will introduce the technology and some of the more important features of the ZED depth cameras, alongside their working principle.

The ZED depth camera is a passive depth ranging tool without an active ranging device because it does not contain an IR laser projector. This depth camera employs a binocular camera to capture 3D scene data, measures the disparity of the objects using a stereo matching algorithm, and finally calculates the depth map according to the sensor parameters [16–20].

The ZED depth sensor is composed of stereo cameras with a video resolution of 2560 × 1440 pixels (2K) with dual 4 mega-pixel RGB sensors. The two RGB cameras are at a fixed base distance of 12 cm. This base distance allows the ZED camera to generate a depth image up to 20 m (40 m is the maximum distance in the new updated firmware, according to the Stereolabs datasheet) [16]. The camera contains a USB video device class supported USB 3.0 port backward compatible with the USB 2.0. standard. It should be noted that the ZED 3D depth sensor is optimized for real-time depth calculation using NVidia Compute Unified Device Architecture technology [16]. Therefore, a corresponding graphical processing unit (GPU) and appropriate computer hardware are required to use it [16–18].

The ZED sensor uses wide-angle optical lenses with a FOV of 110°, and it can stream an uncompressed video signal at a rate up to 100 fps in Wide Video Graphics Array (WVGA) format. The depth image is provided with a 32-bit resolution. Hence, the camera gives a very accurate and precise depth image that describes the depth differences, i.e., the different distances from the plane of the camera. Right and left video frames are synchronized and streamed as a single uncompressed video frame in the side-by-side format. Various configurations and capturing parameters of the on-board image signal processor, such as brightness, saturation, resolution, and contrast, can be adjusted through the SDK provided by Stereolabs [16]. Furthermore, ZED devices support several software packages, called “wrappers,” such as ROS, MATLAB, Python, etc. All these software packages allow the modification of different parameters depending on the user requirements, such as the image quality, depth quality, sensing mode, name of topics, quantity of frames per second, etc. [16–20]. Figure 3 shows the ZED depth sensors [16].

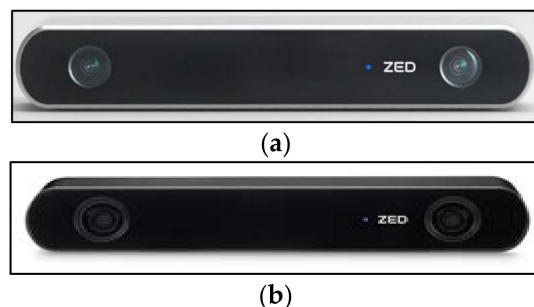
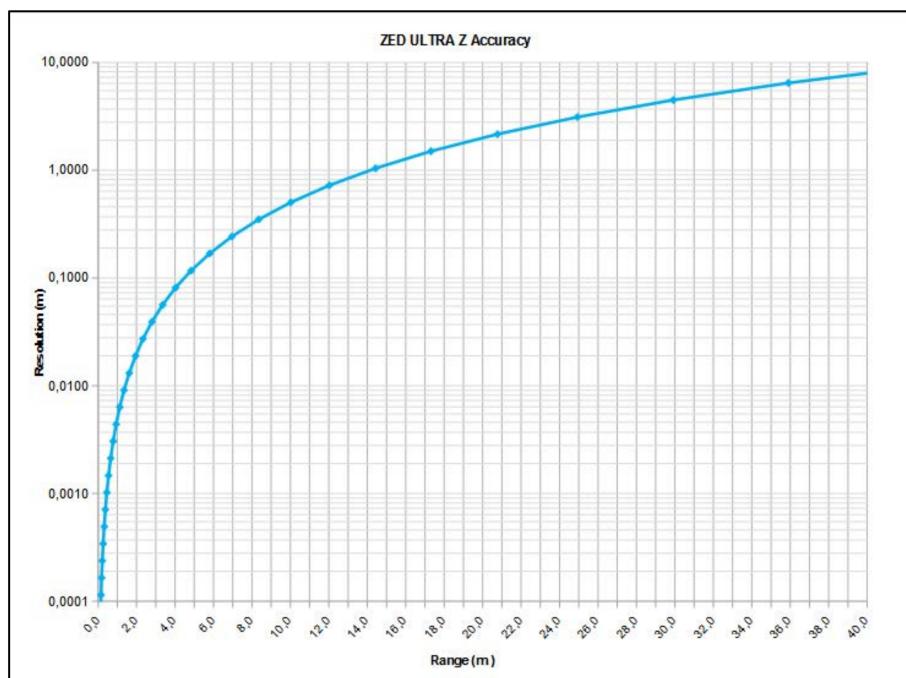
**Figure 3.** (a) ZED and (b) ZED 2i cameras [16].

Figure 4 presents the accuracy graph of the ZED depth sensor depending on the distance of an object from the depth camera. As shown, the depth resolution, i.e., the depth precision, is impaired with increasing distance [16].



**Figure 4.** The accuracy graph of the ZED depth sensor (courtesy of Stereolabs) [16].

Finally, it should be remarked that the ZED depth sensor comes with a unique factory calibration file, which is downloaded automatically. The recommendation is to use the Stereolabs factory settings, but users can also calibrate the ZED sensor with the ZED SDK software package [16–20].

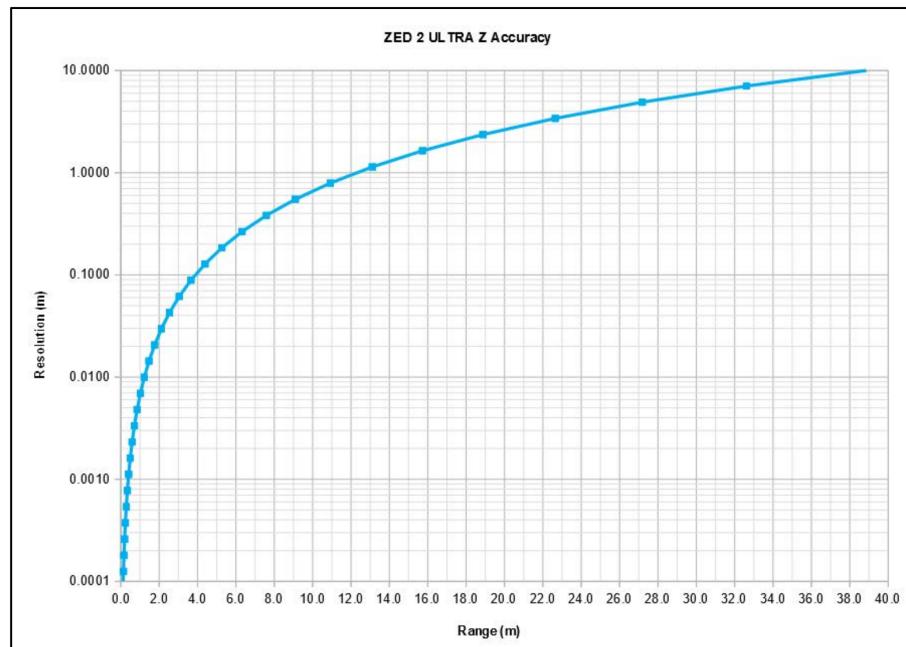
The new ZED 2i depth camera has some similarities and shares some properties with the previous version of the ZED depth camera. However, the new ZED 2i sensor includes several significant improvements.

ZED 2i is the first stereo depth camera that uses artificial neural networks (ANNs) to reproduce human vision, bringing stereo perception to a new level [16]. It contains a neural engine that significantly improves the captured depth image or depth video stream. This ANN is connected to the image DSP, and they contribute jointly to creating the best possible depth map [16]. Furthermore, the ZED 2i camera has a built-in object detection algorithm [16]. This algorithm detects objects with spatial context. It combines artificial intelligence with 3D localization to create next-generation spatial awareness [16]. There is also a built-in skeleton tracking option that uses 18x body key points for the tracking application. The algorithm detects and tracks human body skeletons in real time. The tracking result is displayed via a bounding box, and the algorithm works up to a 10 m range [16].

Next, this ZED 2i camera has an enhanced positional tracking algorithm that is a significant improvement suitable for robotic applications [16–21]. This benefit arises from a wide 120° angle FOV, advanced sensor stack, and thermal calibration for greatly improved positional tracking precision and accuracy [16]. The ZED 2i depth camera has a built-in inertial measurement unit, barometer, temperature sensor, and magnetometer. All these sensors provide extraordinary opportunities for easy and accurate multi-sensor capturing. These sensors are factory calibrated on nine axes with robotic arms [16]. The data rates of the position sensors, i.e., the barometer and magnetometer, are 25 Hz/50 Hz. The built-in motion sensors, accelerometer, and gyroscope contribute significantly to the development of robotic applications since there is no need to install any other sensor except the ZED

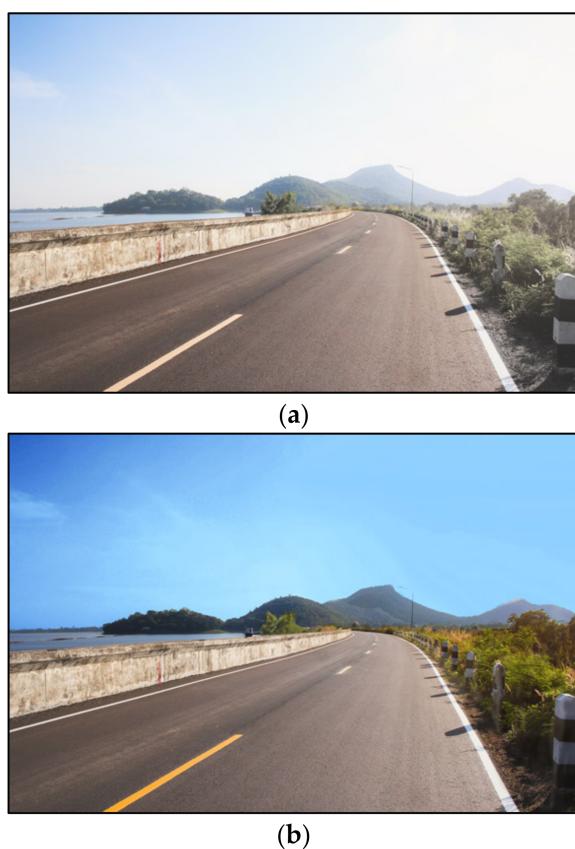
2i camera itself. The data rate of these sensors is 400 Hz. The thermal sensors monitor the temperature and compensate for the drifts caused by heating. In this way, gathered real-time synchronized inertial, elevation, and magnetic field data along with image and depth are collected. These sensors also contribute to accurate and precise depth sensing. The all-aluminum frame reduces the camera heating that induces changes in focal length and motion sensor biases [16]. The case in aluminum allows to better dissipate the internal heat generated by the electronic components reducing the internal temperature of the depth camera. Additionally, the case deformations cannot affect the measure of the depth in any way because the lenses do not move. Furthermore, the software is factory calibrated in order to use the temperature information provided by the internal sensors and modify the data accordingly [16]. All the mentioned motion and positional features indicate that the ZED 2i depth camera is extremely suitable for development of autonomous and industrial robotic applications [21–40].

Figure 5 shows the accuracy graph of the ZED 2i depth sensor, depending on the distance of an object from the depth camera. As shown, the depth resolution, i.e., the depth precision, decreases with increasing distance [16]; however, for instance, at the 1 m range, the accuracy is better than that of the previous ZED, as seen in Figure 4.



**Figure 5.** The accuracy graph of the ZED 2i depth sensor (courtesy of Stereolabs) [16].

One of the most important improvements in ZED 2i is the inclusion of new ultra-sharp eight-element all-glass lenses able to capture video and depth up to a 120° FOV, with optically corrected distortion and a wider  $f/1.8$  aperture, which allows capturing 40% more light [16]. Furthermore, the ZED 2i has an optional feature: the polarizing filter. This built-in polarizing filter gives the highest possible image quality in various outdoors applications. This lens helps to reduce glare and reflections and increases the color depth and quality of the captured images [16]. The effect of the polarizer can be seen in Figure 6.



**Figure 6.** Effect of the polarizer on a scene: (a) without polarizer, (b) with polarizer (courtesy of Stereolabs) [16].

The ZED 2i stereo camera also has two lens options. It is possible to choose between a 2.1 mm lens for a wide FOV or a 4 mm lens for increased depth and image quality at long range, according to the manufacturer [16]. However, the 4 mm lens option entails delays in delivery [16]. These are principal features related to all depth cameras since the lenses, aperture, and light significantly affect the image quality of any camera, not only the depth camera. These features enable high-quality depth images to be obtained, upon which promising robotic vision applications can be developed [41–59]. According to the Stereolabs [16], the main features of ZED cameras are listed in Table 2.

**Table 2.** Features of ZED and ZED 2i cameras [16].

Features	ZED	ZED 2i
Size and weight	Dimensions: 175 × 30 × 33 mm Weight: 159 g	Dimensions: 175 × 30 × 33 mm Weight: 166 g
Depth	Range: 1–20 m Format: 32 bits Baseline: 120 mm	Range: 0.3–20 m Format: 32 bits Baseline: 120 mm
Image sensors	Size: 1/3" Format: 16:9 Pixel Size: 2 μm pixels	Size: 1/3" Format: 16:9 Pixel Size: 2 μm pixels
Lens	Field of View: 110° Six-element all-glass dual lens $f/2.0$ aperture	Field of View: 120° Wide-angle eight-element all-glass dual lens with optically corrected distortion $f/1.8$ aperture

**Table 2.** Cont.

Features	ZED	ZED 2i
Individual image and depth resolution in pixels	HD2K: 2208 × 1242 (15 fps) HD1080: 1920 × 1080 (30, 15 fps) HD720: 1280 × 720 (60, 30, 15 fps) WVGA: 672 × 376 (100, 60, 30, 15 fps)	HD2K: 2208 × 1242 (15 fps) HD1080: 1920 × 1080 (30, 15 fps) HD720: 1280 × 720 (60, 30, 15 fps) WVGA: 672 × 376 (100, 60, 30, 15 fps)
Connectivity and working temperature	USB 3.0 (5 V/380 mA) 0 °C to +45 °C	USB 3.0 (5 V/380 mA) −10 °C to +45 °C
SDK System minimal requirements	Windows or Linux Dual-core 2.3 GHz CPU 4 GB RAM Nvidia GPU with compute capability > 3.0	Windows or Linux Dual-core 2.3 GHz CPU 4 GB RAM Nvidia GPU with compute capability > 3.0
Additional sensors	-	Accelerometer Gyroscope Barometer Magnetometer Temperature sensor
Software enhancements	-	Depth Perception with Neural Engine Built-in Object Detection

Furthermore, the ZED 2i has an Ingress Protection 66 (IP66)-rated enclosure that is highly resistant to humidity, water, and dust, since the ZED 2i depth sensor is designed for outdoor applications and challenging industrial and agricultural environments, etc. [16], while the ZED camera, such as the D415 and D435 RealSense cameras, is mainly designed for indoor applications [7]. Moreover, the ZED 2i has multiple mounting options and a flat bottom, and it can be easily integrated into any system and environment [16].

Since the ZED 2i can be cloud-connected, there is an option to monitor and control the camera remotely. Using a dedicated cloud platform, capturing and analyzing the 3D data of the depth image is possible from anywhere in the world [16]. It is also possible to monitor live video streams, remotely control the cameras, deploy applications, and collect data.

Finally, it should be mentioned that there is also a ZED two-depth camera provided by Stereolabs [16]. The main features and the accuracy graph of the ZED 2i are the same as those of the ZED 2 as the internal sensor stack and the lens configuration of the two sensors are the same [16]. The only differences are the external enclosure for the ZED 2i, which is IP66 and hence more robust, and the option of a built-in polarizing filter [16].

The essential features of ZED and ZED 2i depth devices are summarized and compared in Table 2 [16].

#### 4. Experiments and Results

In this section, the experiments and their outcomes are explained in detail. Numerous measurements in various situations of different scenes and objects were made with the described four depth cameras, and several results are presented here. The scenes are selected in such a way to present many different situations as possible, with many different objects as possible to cover the aspects of possible obstacle avoidance applications and the aspects of possible applications where high accuracy is needed for object detection. Further, the available depth sensors are using different parameters that are discussed in this paper, and all the parameters of used depth cameras are set automatically by the cameras themselves and its software. In that sense, the goal of experiments is to investigate the resulting depth maps and determine the abilities of the depth sensors in particular capturing situations. The fair comparison of the results will be presented, since the depth maps are very informative as the results of the depth measurements, and they provide the best assessment of the capabilities of available depth cameras themselves, specially from the practical point [1–22]. In all the examples, the position and the capturing conditions

such as the illumination, angle, shadows, etc., of all the depth sensors were the same for a particular example in order that the results could be compared appropriately. All the depth maps in examples are displayed in grayscale, since with the ZED software only the grayscale depth maps can be displayed [16]. The RealSense software can display depth maps in various color and grayscale modes [7], but for a better comparison all the depth maps are presented in grayscale in the paper.

Furthermore, it should be noted that all the examples are provided without any post-processing methods, since the goal was to examine, analyze, and present the capabilities of the depth sensors themselves. All the examples were recorded with the built-in options provided by the depth cameras and their recording software. Both the RealSense and the ZED software support the so-called fill option of the depth map, where the objects in the depth image are smoothed and blurred. This option was suitable for examples where the overall visual experience was preferred, such as panoramas. The opposite of this fill option is the standard high accuracy preset, where the edges of the objects in the depth map are more pronounced in order, such that the finer details in the depth image can be efficiently distinguished. This option is mandatory in applications where high precision and accuracy are the main demands. The effects of these options will be presented later in this section.

It should also be mentioned that the experiments were conducted with two kinds of hardware configurations: with an older, modest quality laptop with an on-board graphic card and a new desktop computer with an NVidia GPU. The goal of using both computers in experiments was to explore the capabilities of the depth sensors depending on the quality of the computer equipment to which the depth cameras were connected. In practice, it can happen that an application does not require state-of-the-art computer equipment or a high-quality depth sensor, so it was convenient to check the capabilities of the depth cameras in this case, as well. The latter case primarily refers to applications that do not require high accuracy and precision of depth measuring of the scene, i.e., applications where the depth map of the environment will be used for orientation and navigation in space when moving the robot or a vehicle. Therefore, the presented examples cover this issue as well. Figure 7 shows the equipment utilized in the experiments.



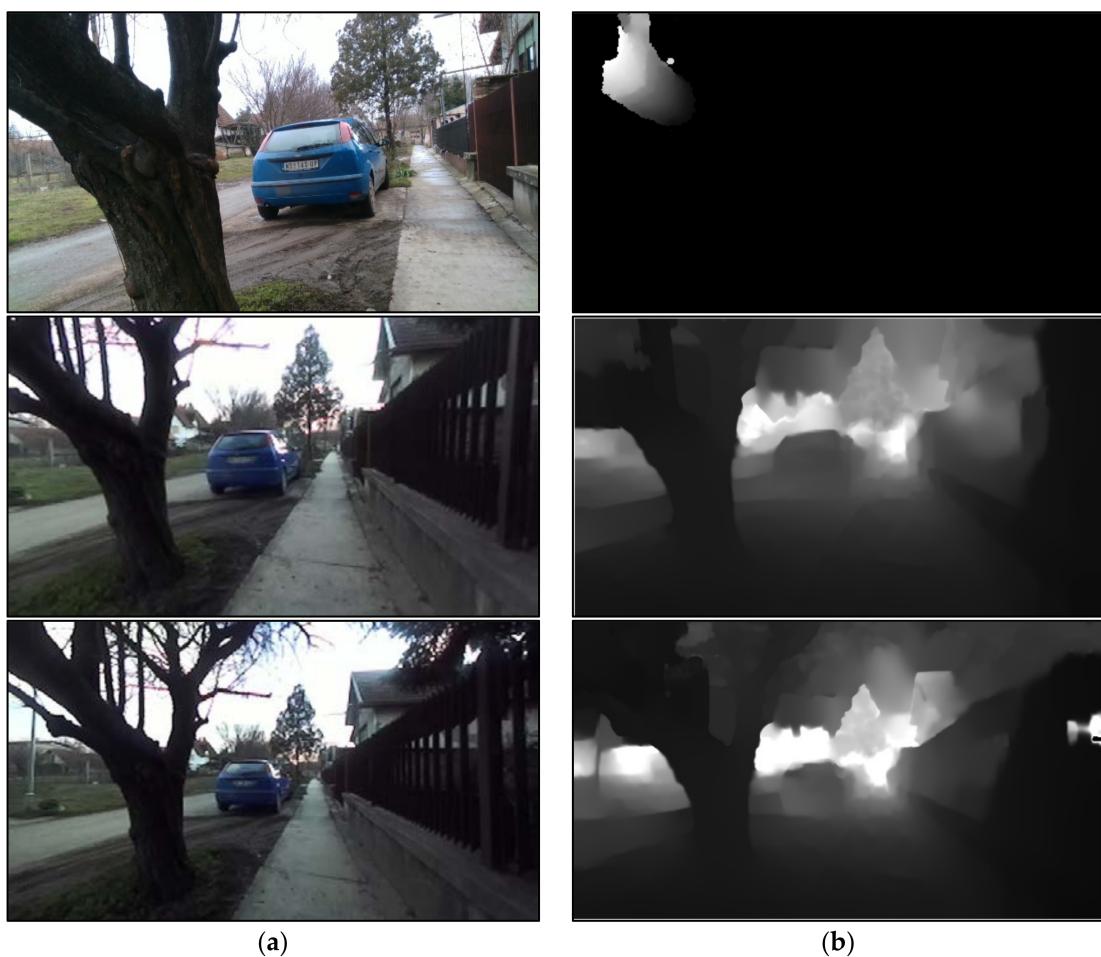
**Figure 7.** The available depth sensors and equipment.

Figure 8 shows color street photos in (a) column and their corresponding depth maps in (b) column, one below another, as recorded with the D435, D415, ZED, and ZED 2i depth cameras, respectively. The first row is the result from the D435 camera, the second row is the result from the D415 camera, the third row is the result from the ZED camera, and the fourth row is the result from the ZED 2i camera. It should be noted that in all further examples the measurement results from the depth sensors are displayed in the same way. Furthermore, the (a) column shows the color images recorded from the left cameras of all the depth sensors, while the (b) column displays the corresponding depth images in all

the presented examples in this section. It is important to highlight that the depth images are always aligned with the color images from the left cameras for all the depth sensors, and this is the reason the left camera's color images are used in all the examples. The color image from the left camera and the depth image share the same set of XYZ coordinates for all depth sensors. It should also be noted that the vertical black, thin areas in the depth images are caused by the distance between the cameras in the depth sensors, and they are called the "dead zones". In this zone, the depth cameras cannot detect anything in depth [7,16]. This is the effect generated by the parallax [16]. Parallax effect is a difference or displacement in the apparent position of an object viewed along two different lines of sight, and is measured by the angle or semi-angle of inclination among those two lines. Hence, the two-camera sensors, the left and the right sensors, look at the scene from a horizontal distance of few centimeters, so it can happen that part of the scene, objects are not visible by both the sensors [7,16]. Accordingly, the first example in Figure 8 shows the results recorded with the RealSense D435 depth camera. The depth camera was set on the sidewalk, and the illumination was satisfactory, without any sunrays, shadows, and non-specific artifacts. The tree was at a distance of about 3 m from the camera, and the car was about 7 m away from the camera. The fence was 0.5 m away from the depth camera, and it moved farther from the depth camera itself with the growing distance. Other trees, houses, road, etc., can be seen in the photo, but at greater distances. The recording conditions and depth sensor positions were the same for all depth sensors. As noticed in the first example in (a) column, there is a rich content in the captured image. In column (b) the corresponding depth map is shown. The depth map was generated with the so-called default preset in the RealSense Viewer [7], which provides the filling option for the depth image. As noted, this option ensures an overall visual experience, and the depth image is pleasant to view. The nearest tree, the nearest part of the fence, and the nearest part of the ground are clearly visible, but on the other hand, the car is very dark and invisible in the depth map. This can be considered a good result, since the D435 depth camera and all the other depth cameras are mainly designed for indoor use [7,16,17]. Based on this result, we concluded that the D435 depth sensor yielded a satisfactory depth image of the panorama, and this depth image could be used for applications where the robot orientation and spatial navigation are required. Finally, this measurement result was generated with a modest quality laptop, where the resolution of the color image and the depth map was  $1280 \times 720$  pixels. The same resolution was used in the RealSense D415 depth camera measurement as well as the same default preset.



**Figure 8. Cont.**



**Figure 8.** Examples of panorama depth measurement: (a) color images from the left camera, (b) corresponding depth maps.

The second example shows the result of measurement with the D415 depth camera. It can be noticed that the scene is a little narrower because of the built-in optics in the depth sensor; hence, the recorded color image in the first column contains fewer details. Unfortunately, the depth map is almost invisible, and only a small part of the scene is present near the tree area. The reason for this lies in the architecture and properties of the D415 camera itself, since the D415 depth sensor is intended for high accuracy measurement of small indoor distances [7]. Therefore, the obtained results were not a surprise, and any useful visible result in the depth image would have been a gain in this case.

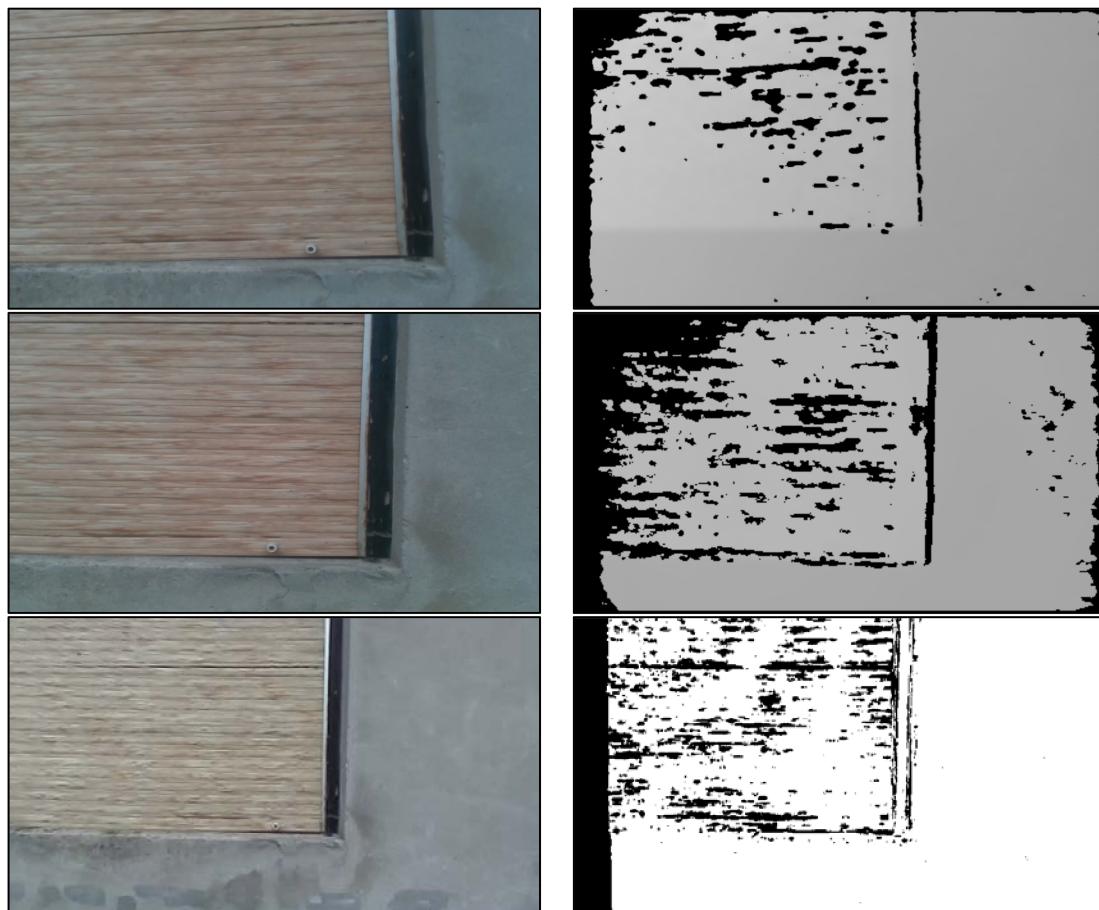
The third example in Figure 8 was recorded with the ZED depth sensor. The color image on the left contains much more detail. The reason for this wide view was the wide FOV of both ZED depth sensors. Moreover, since the modest quality laptop without the GPU was used for the recording, the maximal supported resolution was only  $672 \times 376$  pixels. Despite the poor resolution, the recorded scene is of good quality. The recording was done with the ZED Explorer [16], and the images were saved in \*.svo [16] files. Later, using the desktop computer and the NVidia GPU, the depth maps were generated via the ZED Depth Viewer Tool [16]. The reason for this was the absence of a high-quality GPU on the laptop as required by the ZED depth cameras for depth calculation [16,17]. Unfortunately, the original resolution of the saved images could not be changed during the depth determination; hence, the depth measurement was done with a resolution of  $672 \times 376$  pixels. The corresponding depth map is shown on the right.

Regardless of the stated quality losses, the generated depth image is of very good quality and could be very useful in applications where the robot needs to orient itself in

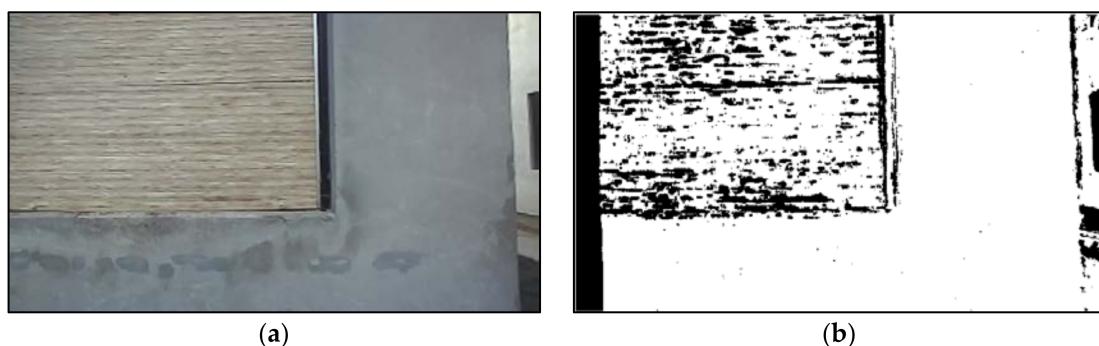
space. In the obtained depth map, not only the tree and the fence are correctly visible, but the car is visible, too, along with the far trees and the houses. Therefore, we concluded that the ZED depth sensor can generate a satisfactory depth map, even with the degraded performance. During the depth calculation, the so-called ultra-quality depth determination [16] was used with the filling option in order to generate a pleasing depth image for an overall visual experience.

Finally, the fourth example shows the result generated with the ZED 2i depth sensor. All the settings in the ZED software [16] were the same for the ZED 2i camera as for the ZED camera. Since the ZED 2i has wider optics and FOV, the color image is the most detailed of all the examples in Figure 8. This is to be expected regarding the features of all the depth sensors in this research. The determined depth image is displayed in the right column and as can be noticed, it contains more details because of the wider angle of view. Additionally, the quality of the depth map is almost the same as that of the ZED camera, and possibly the details are a little sharper. Hence, this depth camera can be very useful in applications where the robot's spatial orientation is mandatory.

Figure 9 presents the example of window frame depth measurement on a facade. This experiment reflected the necessity of accurately determining the depth for industrial robotic applications, such as painting robots [17,20]. The goal in this and similar applications was to separate the background from useful objects or surfaces. In order to achieve this, it was very important to clearly distinguish the edges of objects in depth in the depth images. The depth sensors in all measurements were set 1 m away from the wall. The depth difference between the wall and window was about 8 cm.



**Figure 9.** *Cont.*



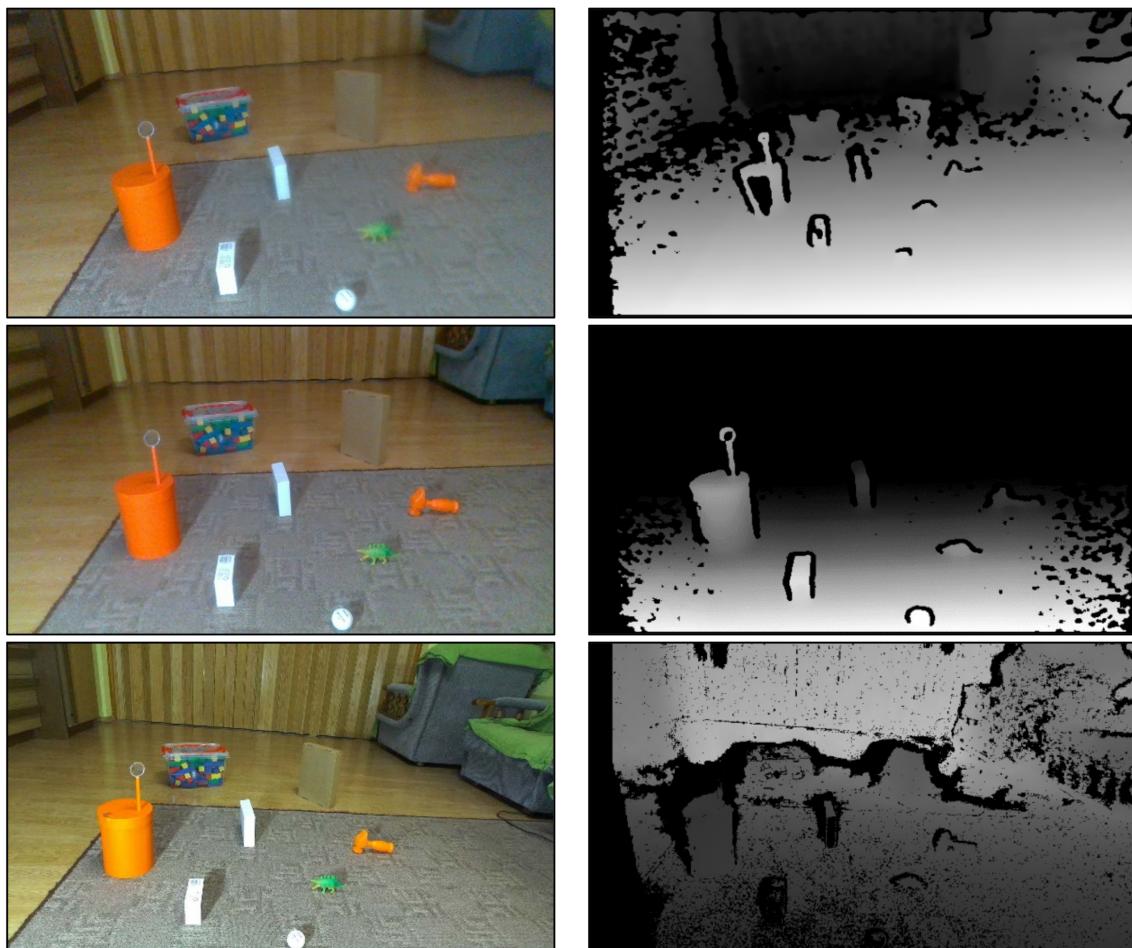
**Figure 9.** Examples of window depth measurement on a facade: (a) color images from the left camera, (b) corresponding depth maps.

In the first two examples in Figure 9 the resolution of the D435 and D415 sensors was  $1280 \times 720$  pixels, and this resolution was used in all the examples in this paper. However, since the accuracy and precision were very important, the high accuracy preset [7] was used in these measurements, since the goal was to distinguish the wall from the window. As we see in first two color images, they are almost of the same quality, without any significant differences. However, the depth maps are very different. It can be seen that in the first example, the depth map recorded with the D435, the window edges are not clearly visible, especially the bottom region. This result will also affect the post-processing, since it would be very difficult to distinguish the window frame in order to separate it from the wall. On the other hand, the second depth image in Figure 9 shows the result generated by the D415 depth camera, where the window frame is clearly visible and distinguishable. Later, this result can be post-processed successfully to separate the wall from the window. However, observing the depth images of the D435 and D415 cameras, one can notice the noise and very poor shade distinction of the surfaces of the wall and the window. The reason for this was the poorer ability of the sensors themselves that are built into RealSense cameras, compared to those in the ZED cameras.

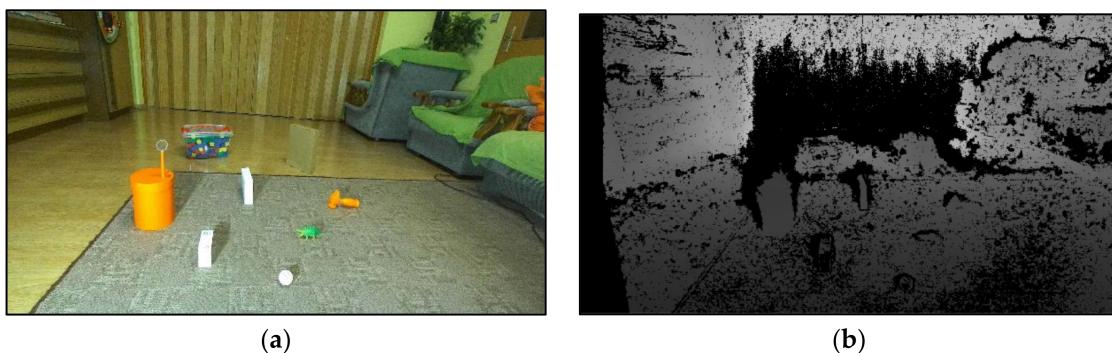
The third and fourth image examples show the results obtained by the ZED and ZED 2i depth cameras. In both examples where the ZED depth sensors were utilized, the image resolution was  $672 \times 376$  pixels, the depth calculation setup was in ultra-quality, but now with the standard preset [16] recommended for high depth accuracy measurements. The wider FOV in both color image samples should first be noticed. Because of the wide FOV in the fourth image obtained from the ZED 2i, part of adjacent building can also be seen in the right side of the image. Therefore, the ZED 2i would be very useful for facade painting application via robots because from one position it can capture a very large wall surface in the image vision stage. Furthermore, in both depth maps the window frame is clearly separated. These results are very suitable for additional image processing in order to extract the wall area intended for painting [20]. However, we noticed that the measured window edges are more preserved in the fourth example, namely, in the depth map measured by the ZED 2i. The reason for this was probably the better lens optics of the ZED 2i camera compared to the previous version, the ZED camera. In the depth map generated with the ZED 2i, the edge of the building is visibly preserved relative to the background. In this example, it was expected that the window should have a well-defined distance determined by depth cameras, and gray level different with respect to the wall. On the contrast, the window area presents many horizontal dark lines and its gray level is the same as of the wall. According to depth camera manufacturers, the texture of the window is not very informative, and the wall texture too, and this problem results the depth maps seen in Figure 9b [16]. Additionally, the window presents many black pixels that correspond to zones with no depth information from the same reason. The final conclusion based on the examples in Figure 9 was that the ZED depth cameras yielded much better accuracy and depth maps than the RealSense depth cameras in measurements where accuracy are

mandatory. Based on the images in Figure 9 the difference in the quality of the equipment capabilities is obvious for possible painting robot application.

The experiment presented in Figure 10 depicts a homemade polygon for robotic application where the main goal is obstacle avoidance. Objects of various dimensions and shapes were randomly arranged in the room at different distances from the depth camera position. The lighting of the room came from a common chandelier with five LED bulbs with so-called natural yellow light. Some of the objects were very small, such as a table tennis ball, a small dinosaur figure, and a magnifying glass placed on a big orange box. The dimensions of the toy hammer, the white boxes, the brown box, and the box with wooden block were bigger. This polygon could serve as a movement testing polygon for a small robot. Naturally, this setup could be expanded with more objects at greater distances from each other in a room of much larger dimensions for larger robots. From the color images, it is obvious that a depth map with great accuracy was needed, since there were small objects in the scene that needed to be avoided. The closest object was about 1 m away from the camera, and the farthest object about 3 m away. Hence, the previously mentioned presets for high accuracy depth measurements were chosen in all depth sensors [7,16].



**Figure 10.** *Cont.*



**Figure 10.** Examples of spatial scene depth measurement in a room: (a) color images from the left camera, (b) corresponding depth maps.

The first example in Figure 10 shows the images captured with the D435 depth camera. The resolution of the depth sensor was  $1280 \times 720$  pixels, as was the resolution of the D415 depth camera in the second example in this figure. As can be seen, the color image is a little blurred because of the illumination, but the image can be considered as satisfactory. The corresponding depth image is shown in the right-hand column. The depth sensor detected the object very superficially, especially the objects with small dimensions. This was an expected result, since the D435 depth camera is not designed for applications where high accuracy is mandatory. However, the objects are detectable in the depth map, and the depth map could be used for an obstacle avoidance application.

The second result presents the measurement obtained with the D415 depth sensor. The quality of the color image is similar, possibly less blurred compared to the color image captured with the D435 camera. Still, the corresponding depth map in the right-hand column is much better, and the objects are more visible in the depth image. The drawback is that the farthest objects were not detected with the D415 sensor. This was because the D415 sensor cannot detect far objects, as was shown in previous examples. However, this problem would not affect the possible robot movement application, since the robot would receive the depth maps correctly, only from shorter distances, and it would process them.

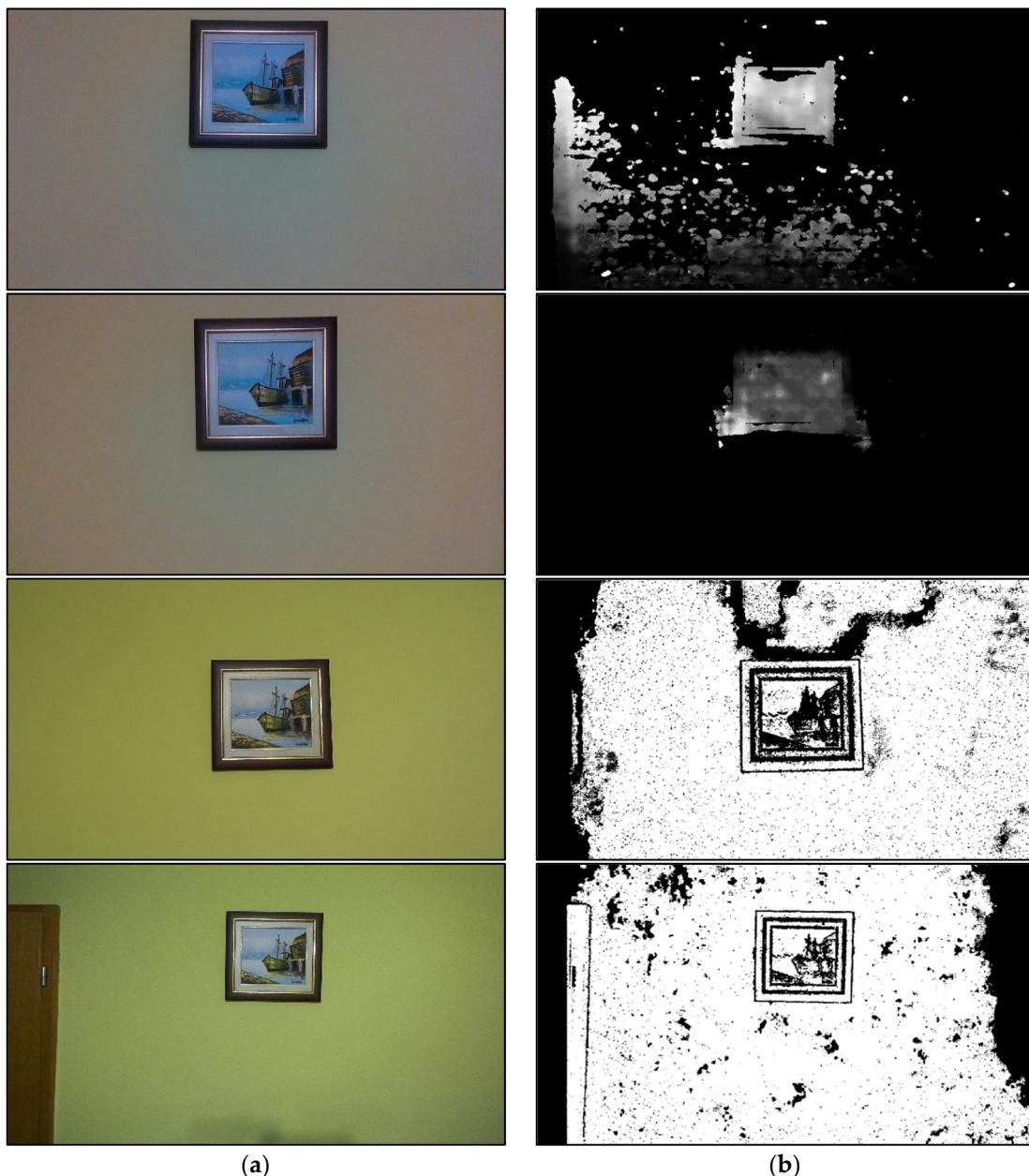
The third and fourth examples show the results captured with ZED and ZED 2i depth sensors. The main difference compared to previous examples is the resolution. In this and in next examples, the resolution of the ZED cameras was  $2208 \times 1242$  pixels for both the color and depth images. The depth calculation mode remained the standard ultra-quality [16]. Again, because of the large FOV, the images contain more details, especially the image captured with the ZED 2i in the fourth row in Figure 10. The recorded color images are high quality, and they are very sharp. In the third example obtained from the ZED depth camera, all the objects in the depth map are visible and detectable; however, the image is not very pleasing to view because of the many details in it. The depth image is considered accurate and precise; hence, it could be used for robot movement applications.

In Figure 10, the fourth example in the second column shows the depth image obtained with the ZED 2i. The yielded depth map contains all the important objects and the large background from the color image. Since the ZED 2i FOV is very large, the depth image has many details, and some of them are not significant for the possible robot movement application. This is considered to be a disadvantage in some situations. However, the preferred objects are accurately recognizable in the depth map. Based on the examples presented in Figure 10, all the depth sensors could be used in robot movement applications for obstacle avoidance where the objects are not very small, since all the sensors yielded satisfactory depth maps. If the objects intended for avoidance are of very small dimensions, the ZED sensors could be the solution, since these sensors can measure and detect small depth differences in practice.

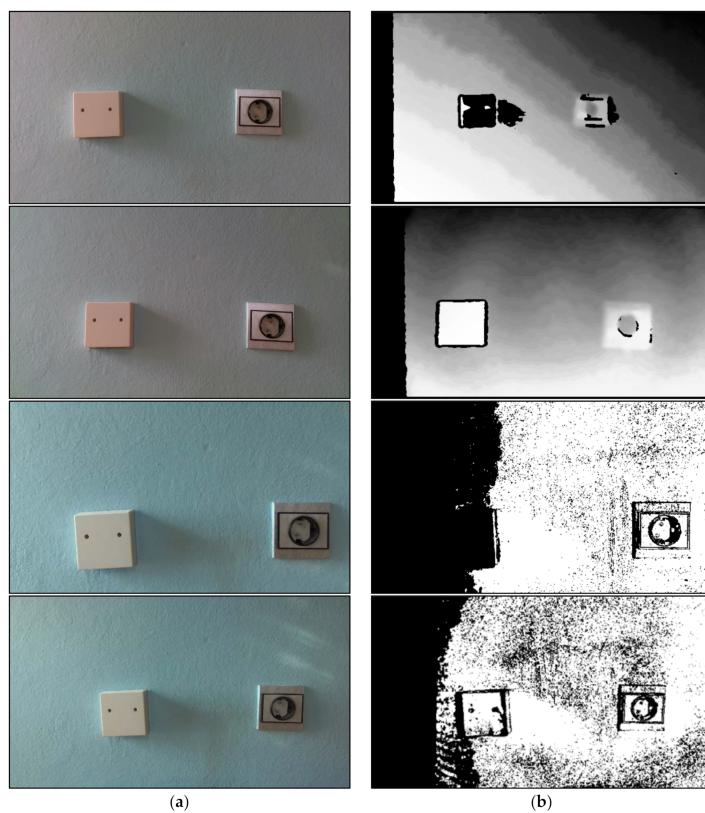
The next three examples were experiments where high measurement accuracy and precision were required. Medical and industrial robotic applications require very accurate

and precise depth maps, and the goal of the following experiments was to investigate the capabilities of depth sensors in such examples and setups. In the examples in Figures 11–13 the high accuracy presets were used with resolutions of  $1280 \times 720$  pixels for RealSense depth cameras and  $2208 \times 1242$  pixels for ZED depth cameras. The RealSense Viewer tool or the RealSense cameras sometimes crashed when the resolution was set to a maximal  $1920 \times 1080$  pixels, and this is the reason why these results were excluded from this paper.

In Figure 11 the depth measurement of a painting is presented, where the painting was the object intended for detection via depth map. The painting frame was about 2 cm thin against the wall area, and the wall was similarly yellow-colored. The height and width of the painting were not significant since the depth camera calculates the depth information in order to generate the depth map. The depth sensors were about 1.2 m away from the wall in all measurements. The lighting of the room was derived from the same common chandelier with five LED bulbs.

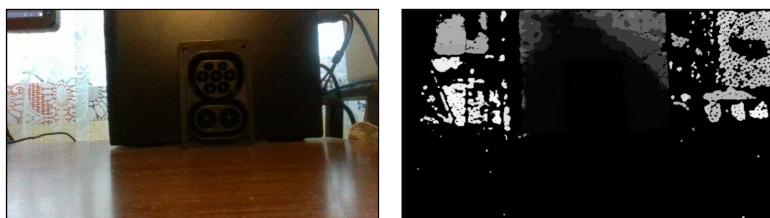


**Figure 11.** Examples of thin painting depth measurement: (a) color images from the left camera, (b) corresponding depth maps.

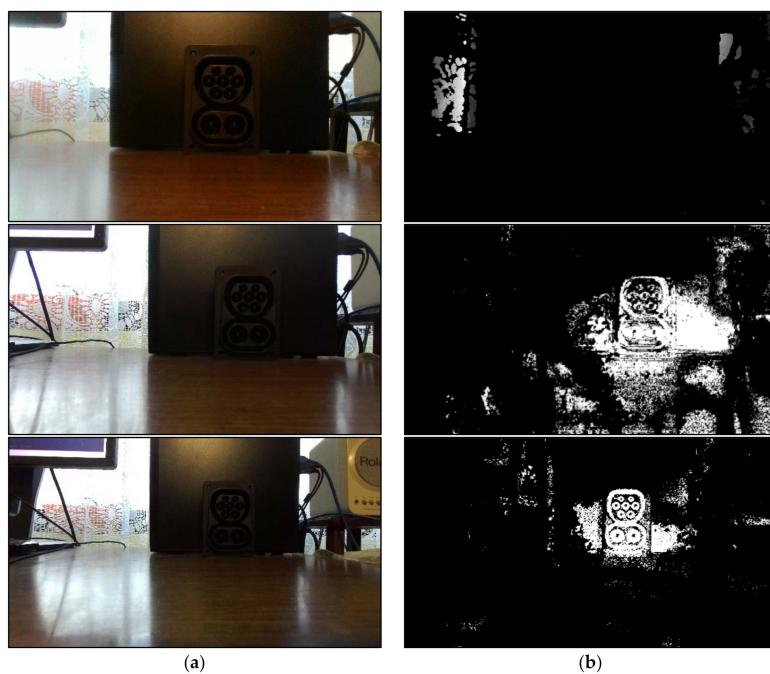


**Figure 12.** Examples of AC socket depth measurement: (a) color images from the left camera, (b) corresponding depth maps.

The first example shows the results yielded with the D435 depth camera. The yellow color of the wall is not overly pronounced, since the image sensors of the camera cannot sense a better color image. Still, the color image is good enough, and all the details are recognizable. The corresponding depth image in the right-hand column is considered as good, since the painting frame is visible, but the painting area is damaged a little and not clearly visible. This shortcoming could cause problems for post-processing algorithms when the goal is the detection and extraction of the painting. There are also some noise and artifacts remaining in the background on the left side and under the painting in the depth map. Hopefully, these artifacts will not affect the detection algorithm for the desired object from the depth image.



**Figure 13. Cont.**



**Figure 13.** Examples of very dark electric car charger socket depth measurement: (a) color images from the left camera, (b) corresponding depth maps.

The second experiment presents the images obtained from the D415 depth sensor. In this example, the color image is much better, since the yellow color of the wall is more pronounced, and the painting details are more visible. Furthermore, the corresponding depth image in the (b) column is more accurate and precise, and the painting frame is detected very well, without the surrounding artifacts and noise. This result was expected, since the D415 depth camera is designed for applications where high accuracy is the main demand, compared to the D435 depth camera where the visual experience is the goal [7]. This depth map is considered as very good, and it could be very usable in robotic vision applications.

The third example shows the measurements obtained with the ZED depth camera. The color image is high quality, and the yellow color of the wall as well as all the details in the image are clearly visible. The corresponding depth map is shown next to the color image, and it is very accurate. All the details in the painting as well as the painting frame itself are clearly visible. However, most of the wall area is visible, too, but this can most likely be eliminated with post-processing algorithms.

The fourth example in Figure 11 presents the results obtained with the ZED 2i depth sensor. The effect of a wide lens is noticeable, since a larger wall surface was covered with this camera. Again, the colors and details in the depth image are clearly visible, even the door on the left is in the scene. The measured depth map is very accurate. All the details in the painting and its frame are clearly visible, even the edges of the door. Again, part of the background wall area is also visible in the depth map, but it and the door area can be eliminated later with post-processing if needed. Furthermore, it can be noticed the drawing inside the picture in depth maps obtained with ZED cameras. The reason of such results could be because the sides of the picture present minimal occlusions that cannot be matched during the stereo matching step of the stereo algorithm. The black borders are zone where the depth information is not available [16]. Additionally, the boat probably has a homogeneous color that does not generate depth results. The boat on the depth map is indeed black, which means that no depth is detected in that area hence the depth map contains some part of the drawing inside the picture area in black color that acts as a texture [16]. However, the edges of the picture area are clearly and accurately

distinguishable in-depth maps measured with ZED sensors, and this information would be used in post-processing.

Based on experiments presented in Figure 11, we concluded that the ZED and ZED 2i depth sensors yielded accurate and precise depth maps, but in certain robotic applications, the depth maps generated with the D435 and D415 depth sensors could be better used for robotic vision purposes, since the picture area is separated from the background. The reason of better depth maps lies in the fact that the RealSense cameras use IR sensors for depth measurement [7], and this is the key reason of the acceptable picture area in depth maps generated by RealSense cameras. However, in the absence of satisfactory computer hardware, the measurements need to be limited on RealSense depth cameras or equipment with similar requirements, since these cameras do not require high quality GPUs for depth calculation.

In Figure 12 an experiment is presented where the goal was to detect the AC socket in the wall. Near the AC socket there was also a junction box. The AC socket was about 7 mm thin, and the junction box, about 2.2 cm thin. The measurements were conducted in daylight, and the windows were to the left of the scene, as was the illumination. The area around the junction box was brighter, since it was closer to the window compared to the AC socket, and it obtained more light. The depth camera distance from the wall was set to the minimum depth limits. The minimum depth limit of RealSense cameras is around 30 cm, and the minimum depth limit of ZED cameras is around 20 cm; hence, these were the distance positions from the wall in these experiments.

In the first example in Figure 12, the D435 depth camera captured a very good color image, as can be seen on the left. However, again the wall color is not pronounced, since the blue color is very poor in the image. In the right-hand column, the depth map is shown. The depth image is not very clear, the AC socket is very hazy, and the junction box is not visible. Hence, this depth map cannot be used for robotic vision applications. This result was not unexpected, since the D435 is not very accurate and not suitable for applications that demand high accuracy and precision [7].

In the second example, the quality of the color image recorded with the D415 is similar, but the corresponding depth map is significantly better. The frame and edges of the junction box are clearly visible, and the desired AC socket is easily distinguishable. However, the upper part of the AC socket is blurred and hazy. This result can be considered satisfactory for robotic vision applications, since it provides enough information for further image processing. This was also an expected result, since the D415 camera is designed mainly for applications with high accuracy requirements at small distances [7].

In first two examples in Figure 12b, the wall area is at the same distance from the depth camera plain, however the first two depth maps show different gray levels inside the wall area. The reason for this phenomenon lies in the next explanation [7,16]. The gray levels depend on the range of the scale used to render the depth map. If the range is short, the “minimum” depth difference can be visible in object at similar distances. On other hand, if the range of the scale is too large, some objects may not be visible in the image. Additionally, the capturing conditions (illumination, shadows, darkness, etc.) can affect to the visualization of depth maps too. Furthermore, when measuring large areas, such as walls whose planes are at a certain distance from the plane of the depth camera, the results may be scattered, dispersed and different gray levels can occur since, in real-life scenarios, the distances between the wall and depth camera are different in relation to the direct distance of the depth camera from the wall surface itself. Such phenomena are possible due to the imperfections of the depth sensors themselves, and due to the illumination and shadows, and can affect the depth calculation results of the stereo algorithm [16]. Hence, only through experiments the best possible depth map can be achieved [7,16]. There are no universal solutions, and the appropriate measuring conditions must be determined for each application and scene separately [7,16].

The third example shows the images obtained from the ZED depth camera. The color image on the left can be considered as perfect. The generated depth map in the right is

very accurate; the AC socket is clearly visible, including the child protection holes (the round white area in the middle of the AC socket). This result was very good, despite the junction box not being visible (which was not the goal of this experiment). The reason why the junction box is not visible lies in the illumination from the left. The junction box got more light, this light blinded the camera sensors, and the objects illuminated from the left, i.e., the junction box, are not visible in the depth map [16]. This is a common problem with passive stereo cameras because the illumination can cause false detection [16].

The final example in Figure 12 shows the results obtained with the ZED 2i depth camera. The color image is very good, and all the details are clear. There are some light reflections above the AC socket, which can cause some negative effects in the calculated depth map. The depth image is displayed on the right. The obtained depth image is noticeably very precise and accurate, and the AC socket and junction box are emphasized well. Quite likely, the polarizing filter eliminated some illumination and light reflection, and the result is the good depth image. Even the small holes in the box and socket are precisely visible in the depth image.

According to images in Figure 12, the ZED 2i depth camera generated the best depth image of all available depth cameras. The reason for this is primarily its improved optics, which include a polarizing filter [16].

The next example in Figure 13 shows the results obtained in measurements where high precision and accuracy were mandatory. Figure 13 shows a special electric car charging socket, the so-called CCS2 socket. This example is very interesting, since the socket was pure black and very dark. The textures of the socket were not clearly visible. The holes around the electric contacts as well the electric contacts themselves were very dark. The background of the socket was also black. The goal of this experiment was to find and analyze the limit case of measurement, when the details of the object itself are very poorly visible and very dark. Dark objects reflect light less, which results in poorer capabilities of all depth cameras since the textures of dark objects are not visible appropriately, and makes it very difficult to calculate the depth of scenes which contain dark objects and materials [7,16]. The lighting of the room was derived from the same common chandelier with five LED bulbs in this example, and the background light was partly derived from sunrays through the window located behind the black background and socket. The depth sensors were set about 30 cm away from the CCS2 socket in all measurements. It should also be noticed that the sunrays from through the window radiated directly at the depth camera, and there were shadows in the image. The goal of such experiment was to analyze the behavior of depth sensors in real applications with various illumination sources acting simultaneously.

In Figure 13 the first example shows the result of recording with the D435 depth sensor. The color image on the left is a little blurry, but the details are clear. The corresponding depth map is displayed on the right. We see that the CCS2 socket is invisible, and the depth map contains lot of noise and artifacts.

The second example shows the result of capturing with the D415 depth camera. The color image on the left is similar, and the depth map on the right is almost pure black, without any useful content. These results were expected since the RealSense cameras cannot handle the light reflection and the direct illumination [7].

The third example presents the measurements obtained with the ZED depth camera. The color image on the left is very good, and it contains more details owing to the wider FOV and better lens. The yielded depth image is displayed in the right-hand column. The socket is visible in the depth map; however, its region is very noisy and fuzzy, and it cannot be clearly separated from the background. Obviously, the dark surface of the CCS2 socket and the direct illumination affected the image sensors, making it difficult to detect the desired area in the image. Still, this is a much better result compared to the results generated with the RealSense depth sensors.

The final measurement in Figure 13 presents the result obtained with the ZED 2i depth sensor. Again, the wider FOV provided more details in the color image, and the polarizer

lens handled the illumination and reflection to a certain extent. The obtained depth map is displayed on the right. The measurement result is significantly better compared to the depth map generated with the ZED. The CCS2 socket is clearly visible, with less noise and artifacts. The electric contact and the holes are also clearly visible; hence, this depth map could be used for robotic vision application. Apparently, the high-quality lens and polarizing filter provide more features in the ZED 2i than the depth cameras with modest quality lenses.

As it can be noticed in Figure 13, a not good illumination reduces the quality of the visual features degrading the performances of the stereo matching algorithm in ZED cameras. Dark objects and materials do not reflect the light (they are absorbing the light) and so there is no appropriate visual information to process. Since ZED cameras has no active illumination, so the darker is the scene, and less visual information is available. If the light is indeed too strong the effect is dual, but the result is mostly the same, the depth map will be poor [16]. The same problem is with the IR laser in RealSense cameras, since dark objects and materials does not reflect the light suitably, there is no suitable visual information to process even with the laser [7]. Sometimes a good lighting system can help with stereo vision, but the dark object must be illuminated using different angles to create shadows that act as visual textures. Unfortunately, this is very hard to achieve and there is no reliable and unique solution [7,16].

Based on the examples in Figure 13, the ZED 2i depth sensor yielded the most accurate depth map in applications where the objects were not clearly visible because of the dark colors and invisible textures, and in the presence of dim illumination and shadows.

The final example in Figure 14 shows the results of the same CCS2 socket, but in this example the scene is illuminated with two small SMD (Surface-Mount Device) LED (Light Emitting Diode) panels with 20 built-in LED diodes.



**Figure 14.** Examples of very dark electric car charger socket depth measurement: (a) color images from the left camera, (b) corresponding depth maps.

This example is very interesting, since the depth maps in Figure 13 were not obtained well, and now the experiment is expanded with the use of artificial illumination source and the measurements were repeated. The expectation is, if the LED panel will highlight the textures of the socket, then the depth maps will be better. The illumination has been done simultaneously from the left and right sides of depth sensors, where the LED panels were placed in the same plain with depth cameras. In this position, the LED panels will not blind the camera sensors, and this is considered as a very important task in artificial lighting [7,16].

As it can be seen in Figure 14, the use of artificial illumination source illuminates the scene and the textures of the socket with its surrounding are more visible in contrary to the images from Figure 13. Obviously, the LED panels showed in Figure 15 provide satisfactory lighting conditions, because in all the color images the scene is more clearly visible. Additionally, it can be noticed that RealSense cameras capture more colorful RGB images, while ZED cameras suppress the influence of the artificial lighting and their images contain less color. This phenomenon can be attributed to the different optics used by ZED cameras, compared to RealSense cameras. On the other hand, ZED depth sensors provide better depth maps, according to the results in Figure 15. The first example was captured with D435 camera. As it can be seen, the frame of the CCS2 socket is poorly visible in the depth image, unfortunately it is not enough to be able to determine some of its details. The second example shows the results obtained with D415 camera. Here, the connector is not visible at all in the generated depth map. Even the use of lasers could not affect the generation of a quality depth maps with RealSense cameras in this experiment. The third example shows the result measured with ZED depth sensor. In the corresponding depth map some parts of the socket are recognizable and the frame of the socket is visible, however, the small holes in the socket are not visible. The final example presents the depth image obtained with ZED 2i depth sensor. As it can be seen, the socket is clearly recognizable in the depth map with all its parts. Obviously, the built-in polarizing lens compensated the light reflections caused by LED panels, which led to the generation of a quality depth map.



**Figure 15.** The SMD LED panel used for illumination.

Based on the examples in Figure 15, the ZED 2i depth sensor yielded the most accurate and realistic depth map in applications where the objects were illuminated with artificial

illumination. As it can be concluded from example in Figure 15, a quality illumination system can improve the visibility of the scene which leads to better noticeable textures in the RGB image. Thanks to the better visible scene, depth sensors can generate a depth map with greater reliability and quality.

## 5. Conclusions

This review research presented a detailed overview of the D435, D415, ZED, and ZED 2i depth sensors. The principal features of the depth sensors were described and compared. A series of experiments were conducted in various lighting conditions with various objects, and adequate measurement comparisons were made. Based on the experiments, all the analyzed depth cameras were suitable for applications where robot spatial orientation and obstacle avoidance were mandatory in conjunction with image vision. On the other hand, in robotic vision applications where high precision and accuracy were mandatory, the ZED and ZED 2i depth sensors yielded better results. Further, it is shown, that the use of an artificial illumination source can improve the depth measurement in certain applications. Finally, in applications where the required objects were very dark and textureless, along with dim illumination and shadows, the ZED 2i depth sensor generated the most reliable and accurate depth map of the real scene.

## 6. Future Works

In the future, the ZED 2i camera should be acquired and utilized for further research in the field of robotic vision for medical applications, automotive applications, and autonomous industrial robot development.

**Author Contributions:** V.T. drafted the manuscript and conceived and performed the experiments. A.T., Z.V., P.S. and Z.S. checked the test results and suggested the corrections. M.K., J.S. and I.B. supervised the research and contributed to the organization of article. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by projects EFOP-3.6.1-16-2016-00004 and 2020-4.1.1-TKP2020 of University of Pécs, and CPV 731100000-6, CPV 73120000-9 and 2020-1.1.2-PIACI-KFI-2020-00173 of University of Dunaujavaros, co-financed by the European Union.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** The authors would like to thank the editors and the anonymous reviewers for their valuable comments that significantly improved the quality of this paper.

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

## References

1. 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* **2019**, *19*, 489. [[CrossRef](#)] [[PubMed](#)]
2. Hu, J.; Niu, Y.; Wang, Z. *Obstacle Avoidance Methods for Rotor UAVs Using RealSense Camera*; IEEE: Piscataway, NJ, USA, 2017; pp. 7151–7155. [[CrossRef](#)]
3. Giancola, S.; Valenti, M.; Sala, R. A Survey on 3D Cameras: Metrological Comparison of Time-of-Flight, Structured-Light and Active Stereoscopv Technologies. In *SpringerBriefs in Computer Science*; Springer: Berlin/Heidelberg, Germany, 2018; ISSN 2191-5768. [[CrossRef](#)]
4. Keselman, L.; Iselin Woodfill, J.; Grunnet-Jepsen, A.; Bhowmik, A. Intel RealSense Stereoscopic Depth Cameras. *arXiv* **2017**, arXiv:1705.05548.
5. Lagendijk, R.L.; Franich, R.E.; Hendriks, E.A. Stereoscopic Image Processing. In *PART I Signals and Systems*; The Work was Supported in Part by the European Union under the RACE-II Project DISTIMA and the ACTS Project PANORAMA; MIT OpenCourseWare: Cambridge, MA, USA, 2002; p. 42.
6. Siena, F.L.; Byrom, B.; Watts, P.; Breedon, P. Utilising the Intel RealSense Camera for Measuring Health Outcomes in Clinical Research. *J. Med. Syst.* **2018**, *42*, 53. [[CrossRef](#)] [[PubMed](#)]

7. Tadic, V. *Intel RealSense D400 Series Product Family Datasheet*; Document Number: 337029-005; New Technologies Group, Intel Corporation: Satan Clara, CA, USA, 2019.
8. Grunnet-Jepsen, A.; Tong, D. *Depth Post-Processing for Intel® RealSense™ D400 Depth Cameras*; Revision 1.0.2; New Technologies Group, Intel Corporation: Satan Clara, CA, USA, 2018.
9. BDTI. *Evaluating Intel's RealSense SDK 2.0 for 3D Computer Vision Using the RealSense D415/D435 Depth Cameras*; Berkeley Design Technology, Inc.: Walnut Creek, CA, USA, 2018.
10. Intel Corporation. *Intel® RealSense™ Camera Depth Testing Methodology*; Revision 1.0; New Technologies Group, Intel Corporation: Satan Clara, CA, USA, 2018.
11. Grunnet-Jepsen, A.; Sweetser, J.N.; Woodfill, J. *Best-Known-Methods for Tuning Intel® RealSense™ D400 Depth Cameras for Best Performance*; Revision 1.9; New Technologies Group, Intel Corporation: Satan Clara, CA, USA, 2018.
12. Grunnet-Jepsen, A.; Winer, P.; Takagi, A.; Sweetser, J.; Zhao, K.; Khuong, T.; Nie, D.; Woodfill, J. *Using the Intel® RealSense™ Depth Cameras D4xx in Multi-Camera Configurations*; Revision 1.1; New Technologies Group, Intel Corporation: Satan Clara, CA, USA, 2018.
13. Intel Corporation. *Intel RealSense Depth Module D400 Series Custom Calibration*; Revision 1.5.0; New Technologies Group, Intel Corporation: Satan Clara, CA, USA, 2019.
14. Grunnet-Jepsen, A.; Sweetser, J.N. *Intel RealSens Depth Cameras for Mobile Phones*; New Technologies Group, Intel Corporation: Satan Clara, CA, USA, 2019.
15. Krejov, P.; Grunnet-Jepsen, A. *Intel RealSense Depth Camera over Ethernet*; New Technologies Group, Intel Corporation: Satan Clara, CA, USA, 2019.
16. ZED Product Portfolio. *Stereolabs Product Portfolio and Specifications*; Revision 1; Stereolabs: Orsay, France, 2022.
17. Tadic, V.; Odry, A.; Burkus, E.; Kecske, I.; Kiraly, Z.; Klincsik, M.; Sari, Z.; Vizvari, Z.; Toth, A.; Odry, P. Painting Path Planning for a Painting Robot with a RealSense Depth Sensor. *Appl. Sci.* **2021**, *11*, 1467. [[CrossRef](#)]
18. Tadic, V.; Odry, A.; Burkus, E.; Kecske, I.; Kiraly, Z.; Odry, P. Edge-preserving Filtering and Fuzzy Image Enhancement in Depth Images Captured by Realsense Cameras in Robotic Applications. *Adv. Electr. Comput. Eng.* **2020**, *20*, 83–92. [[CrossRef](#)]
19. Tadic, V.; Burkus, E.; Odry, A.; Kecske, I.; Kiraly, Z.; Odry, P. Effects of the post-processing on depth value accuracy of the images captured by RealSense cameras. *Contemp. Eng. Sci.* **2020**, *13*, 149–156. [[CrossRef](#)]
20. Tadic, V.; Odry, A.; Burkus, E.; Kecske, I.; Kiraly, Z.; Vizvari, Z.; Toth, A.; Odry, P. Application of the ZED Depth Sensor for Painting Robot Vision System Development. *IEEE Access* **2021**, *9*, 117845–117859. [[CrossRef](#)]
21. Ortiz, L.E.; Cabrera, V.E.; Goncalves, L. Depth Data Error Modeling of the ZED 3D Vision Sensor from Stereolabs. *ELCVIA Electron. Lett. Comput. Vis. Image Anal.* **2018**, *17*, 1–15. [[CrossRef](#)]
22. Jauregui, J.C.; Resendiz, J.R.; Thenozhi, S.; Szalay, T.; Jacso, A.; Takacs, M. Frequency and Time-Frequency Analysis of Cutting Force and Vibration Signals for Tool Condition Monitoring. *IEEE Access* **2018**, *6*, 6400–6410. [[CrossRef](#)]
23. Padilla-Garcia, E.A.; Rodriguez-Angeles, A.; Resendiz, J.R.; Cruz-Villar, C.A. Concurrent Optimization for Selection and Control of AC Servomotors on the Powertrain of Industrial Robots. *IEEE Access* **2018**, *6*, 27923–27938. [[CrossRef](#)]
24. Martínez-Prado, M.A.; Rodríguez-Reséndiz, J.; Gómez-Lozeno, R.-A.; Herrera-Ruiz, G.; Franco-Gasca, L.-A. An FPGA-Based Open Architecture Industrial Robot Controller. *IEEE Access* **2018**, *6*, 13407–13417. [[CrossRef](#)]
25. Flacco, F.; Kröger, T.; De Luca, A.; Khatib, O. A Depth Space Approach to Human-Robot Collision Avoidance. In Proceedings of the 2012 IEEE International Conference on Robotics and Automation, Saint Paul, MN, USA, 14–18 May 2012.
26. Saxena, A.; Chung, S.H.; Ng, A.Y. 3-D Depth Reconstruction from a Single Still Image. *Int. J. Comput. Vis.* **2007**, *76*, 53–69. [[CrossRef](#)]
27. Sterzentsenko, V.; Karakottas, A.; Papachristou, A.; Zioulis, N.; Doumanoglou, A.; Zarpalas, D.; Daras, P. *A Low-Cost, Flexible and Portable Volumetric Capturing System*; IEEE: Piscataway, NJ, USA, 2018; pp. 200–207. [[CrossRef](#)]
28. Carey, N.; Nagpal, R.; Werfel, J. Fast, accurate, small-scale 3D scene capture using a low-cost depth sensor. In Proceedings of the 2017 IEEE Winter Conference on Applications of Computer Vision (WACV), Santa Rosa, CA, USA, 24–31 March 2017; pp. 1268–1276. [[CrossRef](#)]
29. Labb  , M.; Michaud, F. RTAB-Map as an open-source lidar and visual simultaneous localization and mapping library for large-scale and long-term online operation. *J. Field Robot.* **2018**, *36*, 416–446. [[CrossRef](#)]
30. Labb  , M.; Michaud, F. Long-term online multi-session graph-based SPLAM with memory management. *Auton. Robot.* **2017**, *42*, 1133–1150. [[CrossRef](#)]
31. Labb  , M.; Michaud, F. Memory management for real-time appearance-based loop closure detection. In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, San Francisco, CA, USA, 25–30 September 2011. [[CrossRef](#)]
32. Labbe, M.; Michaud, F. *Online Global Loop Closure Detection for Large-Scale Multisession Graph Based Slam*; IEEE: Piscataway, NJ, USA, 2014; pp. 2661–2666. [[CrossRef](#)]
33. Labb  , M.; Michaud, F. Appearance-Based Loop Closure Detection for Online Large-Scale and Long-Term Operation. *IEEE Trans. Robot.* **2013**, *29*, 734–745. [[CrossRef](#)]
34. Fischler, M.A.; Bolles, R.C. Random sample consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Commun. ACM* **1981**, *24*, 381–395. [[CrossRef](#)]
35. Derpanis, K.G. *Overview of the RANSAC Algorithm*; Version 1.2; Computer Science Department, University of Toronto: Toronto, ON, Canada, 2010.

36. Rusu, R.B.; Marton, Z.C.; Blodow, N.; Dolha, M.; Beetz, M. Towards 3D Point cloud based object maps for household environments. *Robot. Auton. Syst.* **2008**, *56*, 927–941. [[CrossRef](#)]
37. Li, X.; Guo, W.; Li, M.; Sun, L. *Combining Two Point Clouds Generated from Depth Camera*; IEEE: Piscataway, NJ, USA, 2013; pp. 2620–2625. [[CrossRef](#)]
38. El-Sayed, E.; Abdel-Kader, R.F.; Nashaat, H.; Marei, M. Plane detection in 3D point cloud using oc-tree-balanced density down-sampling and iterative adaptive plane extraction. *IET Image Process.* **2018**, *12*, 1595–1605. [[CrossRef](#)]
39. Gallo, O.; Manduchi, R.; Rafii, A. CC-RANSAC: Fitting planes in the presence of multiple surfaces in range data. *Pattern Recognit. Lett.* **2011**, *32*, 403–410. [[CrossRef](#)]
40. Mufti, F.; Mahony, R.; Heinzmann, J. *Spatio-Temporal RANSAC for Robust Estimation of Ground Plane in Video Range Images for Automotive Applications*; IEEE: Piscataway, NJ, USA, 2008; pp. 1142–1148. [[CrossRef](#)]
41. Nurunnabi, A.; West, G.; Belton, D. Outlier detection and robust normal-curvature estimation in mobile laser scanning 3D point cloud data. *Pattern Recognit.* **2015**, *48*, 1404–1419. [[CrossRef](#)]
42. Prakash, S.; Kumar, M.V.; Ram, R.S.; Zivkovic, M.; Bacanin, N.; Antonijevic, M. Hybrid GLFIL Enhancement and Encoder Animal Migration Classification for Breast Cancer Detection. *Comput. Syst. Eng.* **2022**, *41*, 735–749. [[CrossRef](#)]
43. Li, Y.; Li, W.; Darwish, W.; Tang, S.; Hu, Y.; Chen, W. Improving Plane Fitting Accuracy with Rigorous Error Models of Structured Light-Based RGB-D Sensors. *Remote Sens.* **2020**, *12*, 320. [[CrossRef](#)]
44. Schwarze, T.; Lauer, M. *Wall Estimation from Stereo Vision in Urban Street Canyons*; IEEE: Piscataway, NJ, USA, 2013; pp. 83–90. [[CrossRef](#)]
45. Xu, M.; Lu, J. Distributed RANSAC for the robust estimation of three-dimensional reconstruction. *IET Comput. Vis.* **2012**, *6*, 324–333. [[CrossRef](#)]
46. Kovacs, L.; Kertesz, G. *Hungarian Traffic Sign Detection and Classification Using Semi-Supervised Learning*; IEEE: Piscataway, NJ, USA, 2021; pp. 000437–000442. [[CrossRef](#)]
47. Zhou, S.; Kang, F.; Li, W.; Kan, J.; Zheng, Y.; He, G. Extracting Diameter at Breast Height with a Handheld Mobile LiDAR System in an Outdoor Environment. *Sensors* **2019**, *19*, 3212. [[CrossRef](#)]
48. Deschaud, J.E.; Goulette, F. A Fast and Accurate Plane Detection Algorithm for Large Noisy Point Clouds Using Filtered Normals and Voxel Growing. In *3DPVT*; Hal Archives-Ouvertes: Paris, France, 2010.
49. Najdataei, H.; Nikolakopoulos, Y.; Gulisano, V.; Papatriantafilou, M. *Continuous and Parallel LiDAR Point-Cloud Clustering*; IEEE: Piscataway, NJ, USA, 2018; pp. 671–684. [[CrossRef](#)]
50. Sproull, R.F. Refinements to nearest-neighbor searching ink-dimensional trees. *Algorithmica* **1991**, *6*, 579–589. [[CrossRef](#)]
51. Tadic, V.; Odry, A.; Kecskes, I.; Burkus, E.; Kiraly, Z.; Odry, P. Application of Intel RealSense Cameras for Depth Image Generation in Robotics. *WSEAS Trans. Comput.* **2019**, *18*, 2224–2872.
52. Aggi, D.; Mazzia, V.; Chiaberge, M. Local Motion Planner for Autonomous Navigation in Vineyards with a RGB-D Camera-Based Algorithm and Deep Learning Synergy. *Machines* **2020**, *8*, 27. [[CrossRef](#)]
53. Yow, K.-C.; Kim, I. General Moving Object Localization from a Single Flying Camera. *Appl. Sci.* **2020**, *10*, 6945. [[CrossRef](#)]
54. Qi, X.; Wang, W.; Liao, Z.; Zhang, X.; Yang, D.; Wei, R. Object Semantic Grid Mapping with 2D LiDAR and RGB-D Camera for Domestic Robot Navigation. *Appl. Sci.* **2020**, *10*, 5782. [[CrossRef](#)]
55. Kang, X.; Li, J.; Fan, X.; Wan, W. Real-Time RGB-D Simultaneous Localization and Mapping Guided by Terrestrial LiDAR Point Cloud for Indoor 3-D Reconstruction and Camera Pose Estimation. *Appl. Sci.* **2019**, *9*, 3264. [[CrossRef](#)]
56. Tadic, V.; Odry, A.; Toth, A.; Vizvari, Z.; Odry, P. Fuzzified Circular Gabor Filter for Circular and Near-Circular Object Detection. *IEEE Access* **2020**, *8*, 96706–96713. [[CrossRef](#)]
57. Odry, Á.; Kecskes, I.; Sarcevic, P.; Vizvari, Z.; Toth, A.; Odry, P. A Novel Fuzzy-Adaptive Extended Kalman Filter for Real-Time Attitude Estimation of Mobile Robots. *Sensors* **2020**, *20*, 803. [[CrossRef](#)] [[PubMed](#)]
58. Chen, Y.; Zhou, W. Hybrid-Attention Network for RGB-D Salient Object Detection. *Appl. Sci.* **2020**, *10*, 5806. [[CrossRef](#)]
59. Shang, D.; Wang, Y.; Yang, Z.; Wang, J.; Liu, Y. Study on Comprehensive Calibration and Image Sieving for Coal-Gangue Separation Parallel Robot. *Appl. Sci.* **2020**, *10*, 7059. [[CrossRef](#)]