COMPARISON OF 3D SENSING DEVICES

Ľubomír URBLÍK*, Jakub BLIZMAN, Erik KAJÁTI, Peter PAPCUN
Department of Cybernetics and Artificial Intelligence, Faculty of Electrical Engineering and Informatics,
Technical University of Košice, Letná 9, 042 00 Košice, Slovak Republic, Tel. +421 55 602 2569,
E-mail: *lubomir.urblik@tuke.sk

ABSTRACT

This paper compares two 3D sensing devices – a stereo camera and a Lidar. Both devices offer different modes, which perform differently, and have strengths and weaknesses. We have conducted a series of tests to assess the usability of these devices in our future research. The first round of tests focused on the interior performance and compared the various modes available. The tests were performed in good and bad conditions, showing best and worst-case scenarios. The second test round was performed outside with natural sunlight and more distant objects. The stereo camera was very susceptible to changes in light intensity and performed worse in the interior tests, but better in the exterior tests. The Lidar struggled with dark surfaces at closer ranges and during the exterior testing due to objects outside its range.

Keywords: *3D sensing, stereo camera, Lidar, depth map*

1. INTRODUCTION

The application of 3D sensing in intelligent environments is a key topic in modern technologies, emphasizing accurate and efficient 3D data capture and processing. With the increasing popularity and availability of various types of sensors, such as Lidar and stereo cameras, the need to create applications that can control these sensors efficiently and exploit their potential to the maximum arises. Despite significant progress in this area, many challenges and opportunities still need to be addressed.

Examples of such applications include self-driving vehicles, which measure distances from other vehicles, pedestrians, or obstacles on the road. This can extend to robot vacuums, which often need to solve similar problems, albeit on a much smaller scale. 3D sensing can also be used for home security – in smart locks for face detection, or in low-light conditions as intrusion detection systems. In industrial automation, such sensors can be used for inventory management, as they can measure the volume of materials in storage bins, quality control, and object manipulation in dynamic environments. Archaeology uses 3D sensing to recreate historical sites and buildings, such as the pyramids, in digital forms and detect places, objects, or ruins difficult to see by the naked eye.

The challenge lies in efficiently capturing, storing, and visualizing 3D image data from different sensors in real-time. Sensors such as Lidar or stereo vision cameras work on different depth capture methods, and each type of sensor has its specific requirements for capturing, processing, visualizing data, and overall integration into the system, which adds to the complexity of the solution.

2. 3D SENSING

3D sensing is the process of capturing information about height, width, and depth. There are many methods of 3D sensing, but this paper will focus on two of them – stereovision and time of flight (TOF).

2.1. Stereovision

Stereovision is the process of reconstructing depth from two images of the same scene taken from different angles. Stereovision cameras can be divided into two categories [1]:

- Passive – the cheapest available 3D sensors, using components available in typical cameras. Various offerings are available, making them accessible and usable in various scenarios. The dependency on visible light is the main drawback of this type of camera, as it loses performance in low-light situations and is unusable without light.

- Active – add an infrared pattern projector, which helps capture a more detailed image of the environment and provides a helpful guide in low or no light situations. Their main drawback is the range of the projector, which makes them unsuitable for long-range applications.

This paper will describe the use of Stereolabs ZED 2i [2], which uses two passive cameras. The images are then combined using feature matching, and with the application of SLAM, the resulting depth map is created [3].

2.2. Time of Flight (TOF)

Lidar sensors use lasers to measure the distance from an object, utilizing a similar technique to ultrasound sensors – the ray is projected onto an object, and the device measures the time it takes for the ray to bounce back into the IR sensor. The object, or scene, is then reconstructed from a large number of measurements into a depth map or a point cloud. These can then be converted into a 3D model [4]. Their main drawback is a lower performance in rain, snow, fog, or dust, as it changes the angle or the distance of the ray. This paper will describe the use of Vzense DCAM550U [5].

2.3. Comparison of used devices

We will focus on comparing these two devices, mainly how they react to different conditions, such as low light, interior or exterior. Both sensors offer different modes, such as quality, performance, and ultra.

Table 1 Device parameters

Device	Resolution	FPS	Range
ZED 2i	4416x1242	15-100	20m
DCAM550U	640x480	15-30	0.35m - 6m

The stereo camera has an advantage regarding the choice of resolution and FPS. The higher the resolution, the more detail can be captured. It also has a longer range, which depends on good-quality lighting. The Lidar has a much smaller range but can work in lower-light scenarios. The parameters are described in Table 1.

3. RELATED WORKS

Wang [6] provides a comparison of Lidars and cameras in the context of autonomous driving. The author explains the challenges and problems arising from each technology and gives examples of tasks suitable for said technology. The technologies are also compared with the capabilities of a human driver in various scenarios, such as low light, poor weather, or distance estimation. The conclusion explains the benefits of using multiple sensors to diminish the deficiencies of each sensor.

Jang et al. [7] performed experiments to evaluate and compare how active sensor attributes affect the resulting depth map in different stereo matching methods. In the window-based method, the authors found that increasing the pattern intensity results in better performance up to a certain point, after which it does not increase any more. The global gain, on the other hand, has no effect on the performance. The same changes occur in global methods, but at a smaller scale. The differences in performance are not as apparent as in the window-based method.

Abdelsalam et al. [8] tested the performance of ZED 2i camera at various resolutions and distances in interior environments. The authors analyzed 80 images, taken from different distances and resolutions. The maximum range is 18 meters, slightly lower than the 20 meters in the data sheet. The range is affected by resolution, with higher resolution offering more range. Their testing shows HD1080 as the optimal setting.

Rustler et al. [9] compared four cameras using a table, a doll, and 10 objects at various distances. The authors highlight the importance of selecting the right tool for the job, as the cameras possess different strengths and weaknesses. ZED 2, a predecessor of ZED 2i, achieved the best mean performance and the highest range.

Garigipati et al. [10] approached this area with a different problem – the combination of SLAM algorithms and sensors best suited for different scenarios. Their testing included eight algorithms and used two Lidars and two stereo cameras. Experiments were performed in the interior and the exterior, with the exterior tests also comparing the positioning of the sensors on the vehicle. The results show the best position of the sensor and the best algorithm to use in a variety of environments and situations.

4. PYTHON APPLICATION

The application's main goal was to provide a way of capturing depth map streams and saving them in a unified format. The app also allows users to change the device parameters (resolution, FPS, etc.) on the fly.

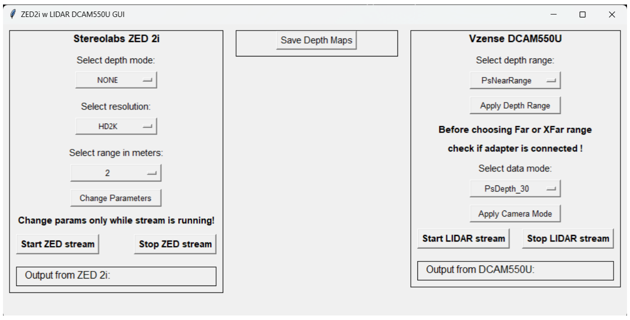

Fig. 1 The GUI of our application

The GUI, shown in Fig. 1, was created using *Tkinter* library [11], which offers predefined tools and widgets, such as buttons, text fields, menus, and more. Another important part of the app is utilizing the *Threading* library [12]. Python runs in a single thread, which prevents us from capturing a video stream from multiple sources simultaneously. By creating multiple threads, our app can run in one thread, the first stream in the second thread, and the second stream in the last thread, ensuring the app runs smoothly.

The stereo camera, Zed 2i, comes with a vendor-supplied Python SDK, which we utilized to control the camera's parameters and capture the data. We then use OpenCV to show the captured data. The Lidar is implemented similarly, with some slight changes due to a different library supplied by the vendor. The largest difference is the need to convert and normalize the captured data before showing the resulting image.

5. TESTING AND EVALUATION

Our testing focused on the performance of selected devices in different scenarios, with different surface types, and with different device parameters. The Zed 2i offers three modes:

- Performance – fluid capture even in worse lighting conditions at the cost of resolution and loss of detail
- Quality – comes with a strong filter, which smooths out sharp edges and surfaces
- Ultra – the mode recommended by the vendor as it offers the highest range and the least loss of detail

The camera offers the following resolutions:

- VGA (1334x376) – the highest framerate with up to 100 FPS and an ultrawide field of view (FOV)

- HD720 (2560x720) – framerate up to 60 FPS and an ultrawide FOV
- HD1080 (3840x1080) – framerate up to 30 FPS and a wide FOV
- HD2K (4416x1242) – the lowest framerate, 15 FPS, and a wide FOV

The resolution is tied to the FOV of the camera, and as can be seen in Fig. 3, there are measurable differences. HD720 and VGA offer the largest FOV, with HD2K coming in second. HD1080 offers the smallest FOV, which can capture more detail.

Fig. 2 The scene used in the first round of testing

The DCAM550U offers four modes, the largest difference being the device's maximum range. The first two modes use just the power delivered by the 5V USB, but the latter two require an external 12-24V power source. The modes are:

- PsNearRange – for close range between 35 cm and 1.5 m
- PsMidRange – for medium range between 50 cm and 2.8 m
- PsFarRange – for far range between 80 cm and 4.4 m
- PsXFarRange – for maximum range between 1.2 m and 6 m

Fig. 3 Difference of FOV by resolution

Unlike the previous device, the Lidar cannot change its resolution, which is fixed at 640x480 (VGA). The framerate is 30 FPS, except in PsXFarRange, which supports 15 FPS only.

The testing was performed in two interior scenes, shown in Fig. 2, and one exterior scene, which will be described later. The testing was done in well-lit conditions, as seen in the images.

5.1. Stereo camera Mode Evaluation

Our testing was performed using the HD1080 resolution, as it offers the highest level of detail with sufficient framerate. The first test compared different available modes, as shown in Fig. 4. Performance captures more objects from a greater distance but with less detail and more noise. Quality mode offers smoother textures but does not capture objects located on the far wall. Ultra mode offers the highest detail but is the most susceptible to changes in light intensity.

Fig. 4 Comparison of Performance, Quality, and Ultra modes on the stereo camera

5.2. Lidar Mode Evaluation

Unlike the stereo camera, which uses grayscale, the Lidar shows the distance using colour. Fig. 5 shows the values used, with red being the closest and purple the furthest.

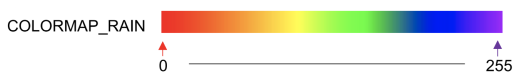

Fig. 5 Colors used by the Lidar

Fig. 6 shows the device modes' results. The distance of the sensor from the far wall was around 5.5 m. From the left figure, it is clear that the *Near* mode is not usable in environment sensing and is more aimed at single-object sensing. *Near* and *Mid* modes both hit the range limit directly in front of the Lidar. Some objects were not detected correctly; for example, the monitor placed in the far left center of the image is shown to be much closer than the house it is placed on. The trash can in the center of the image and some of the objects placed on the shelves were also detected, even though they should be out of range. This is primarily due to the reflectivity of different surfaces and colors. The manufacturer recommends objects with 20 % - 80 % infrared reflectivity at 850 nm wavelength [5]. The *Far* mode detected the entire environment, with some issues with specific objects, but overall is suitable for environment sensing.

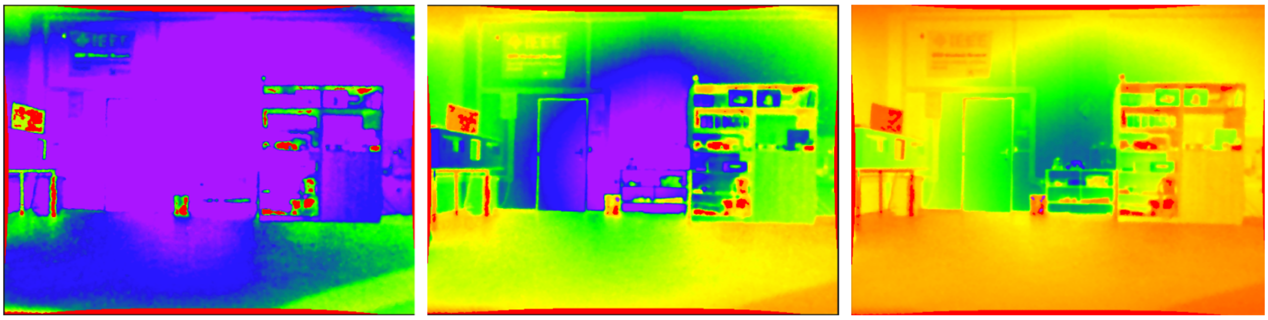

Fig. 6 Comparison of Near, Mid, and Far modes on the Lidar

5.3. Different lighting conditions

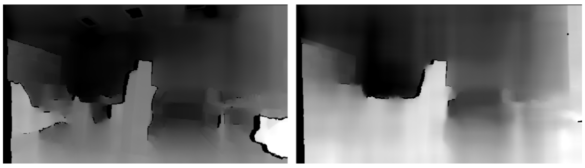

Fig. 7 Stereo camera in a well-lit environment and in dusk

The second round of testing focused on different scenarios, i.e., lighting conditions and how they affect the result, specifically on white and black objects. The stereo camera was tested in a well-lit environment and in the dusk, and used the second scene shown in Fig. 8. As can be seen in Fig. 7, the objects detected in a well-lit environment can be distinguished from one another. This becomes impossible in lower-light situations, with objects blending into one another.

Fig. 8 The scene used in the second round of testing

The Lidar does not face the same problems and performs the same in both scenarios. Nevertheless, it has its drawbacks, namely dark surfaces. As shown in Fig. 9, the device cannot detect a black surface at a distance of 1.5 m and shows it as out of range.

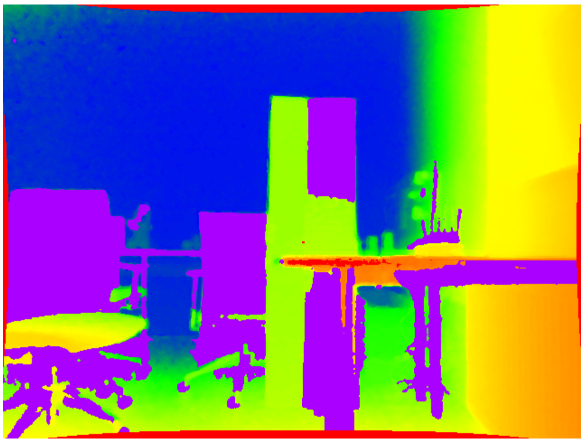

Fig. 9 Comparison of Lidar performance on light and dark surfaces

5.4. Exterior

The first column in Fig. 10 shows two scenes – the first one is aimed at utilizing the maximum range of the Lidar and the second one at the average usable range. Both devices captured both scenes. We used *Ultra* mode with HD1080 resolution on the Zed 2i. These are shown in the center column of Fig. 10 and show a much higher level of detail than the interior scenes. This might be due to the color of the selected scene and/or the natural sunlight. The objects are discernible from one another. The Lidar was set to *XFar* mode, which should have a range of 6 m. The tree on the left side of the first scene is more than 6.5 m away from the sensor. The detail is lost at such a distance, but the device may be usable outside the 6 m range in ideal conditions.

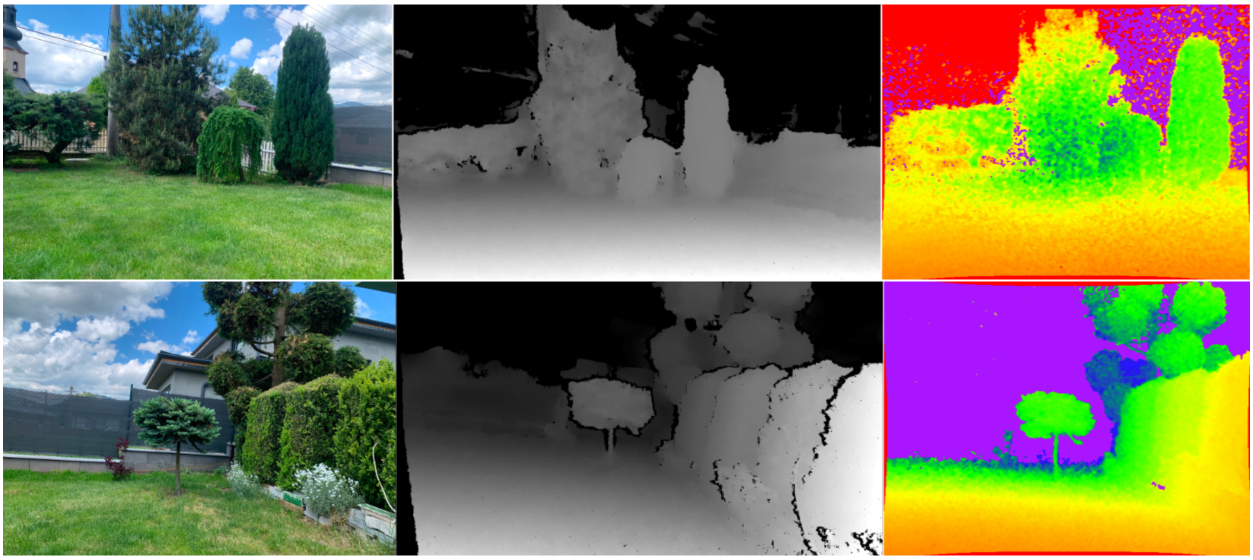

Fig. 10 Comparison of stereo camera and Lidar performance in exterior scenes

Table 2 Hardware load during testing

Device	Mode	Resolution	FPS	GPU	CPU
Together	Ultra/PsXFar	HD720/VGA	100/15	95%	20%
DCAM550U	PsNearRange	VGA	30	3%	5%
	PsMidRange		30	4%	5%
	PsFarRang		30	4%	6%
	PsXFarRange		15	1%	4%
ZED 2i	Performance	VGA	100	31%	11%
	Performance	HD720	60	35%	15%
	Performance	HD1080	30	30%	15%
	Performance	HD2K	15	28%	8%
	Quality	VGA	100	82%	13%
	Quality	HD720	60	89%	15%
	Quality	HD1080	30	90%	14%
	Quality	HD2K	15	43%	9%
	Ultra	VGA	100	85%	13%
	Ultra	HD720	60	96%	15%
	Ultra	HD1080	30	85%	17%
	Ultra	HD2K	15	51%	11%

5.5. Hardware Load

As these devices offer a multitude of modes, with different resolutions and frame rates, the hardware load also changes. The stereo camera, ZED 2i, uses the GPU to process the data, while the Lidar, DCAM550U, uses only the CPU. The values shown in Table 2 were measured during the previously mentioned tests on a device with an Intel i7-8700k, RTX 2080 8GB, and 32GB of memory. The memory is absent from the table, as it did not change in a meaningful way during the tests, always staying at around 28% or 9.2 GB.

The results show a significant difference in the hardware load between the Lidar and the stereo camera, staying at around 5% CPU usage in all modes. ZED 2i shows large differences between modes and resolutions, with the highest resolution, HD2K – approximately 82.23 megapixels per second, being consistently the easiest to run. This is mainly due to the lower framerate. HD720 – approx. 110.59 megapixels per second is the hardest to run in Performance and Ultra modes and ties with HD1080 – approx. 124.416 megapixels per second, in Quality mode. Our testing shows that megapixels per second is not a good indicator of hardware requirements, and the real calculations are more complex than a simple multiplication. HD1080 and VGA have similar GPU usage in Performance and Ultra modes, even though HD1080 has twice the megapixels per second compared to VGA. We also tested the load of both devices running simultaneously, and the results show no difference in GPU usage and a slight increase in CPU usage when compared to running only the ZED 2i.

6. CONCLUSION

This paper focused on comparing two different 3D sensing devices: a stereo camera and a Lidar. We started by developing an application which allowed us to test the devices more efficiently in different scenarios. The devices were tested in interior and exterior environments and various conditions.

Our interior experiments show that the stereo camera depends on good lighting conditions. Lidar was the opposite, where very high and very low light intensities did not make any difference. However, it cannot detect dark surfaces, as the light rays do not reflect correctly.

The exterior tests highlighted the performance of the stereo camera. Natural light proved ideal for this device, as it could easily detect objects in its range and with good detail. The Lidar performed sufficiently, but its lower range and inability to detect dark surfaces make it a worse choice.

Our testing also shows a large difference between the computational requirements of these two devices, with the stereo camera being many times more difficult than the Lidar. These two devices can be used in tandem, resulting in only a slightly higher load than running the stereo camera on its own. By combining the data from both devices, we can obtain more accurate results and mitigate the devices' shortcomings in non-ideal conditions, such as low-light, rain, or a scene with very dark objects.

The retail price for DCAM550U is $369, and for ZED 2i $519. There is no clear winner in this comparison, as both of these devices showed some strengths and some weaknesses. The Lidar is better for shorter ranges or where the surfaces reflect the rays properly. The stereo camera performs better for longer ranges and good lighting conditions, capturing more detail than the Lidar. As mentioned before, we believe the best way forward is to combine data from both devices.

As for the possible future direction of our work, the stereo camera offers another mode, called *neural*, which uses artificial intelligence to enhance the images. A testing and comparison of this mode may yield interesting results. Another possible direction is to create a full 3D model of an environment, i.e., our laboratory, using the tested devices. We also plan on combining the data from both devices to see if they can truly supplement one another.

ACKNOWLEDGEMENT

This publication was supported by EDEN: EDge-Enabled intelligeNt systems, project number VEGA 1/0480/22.

REFERENCES

[1] RODNITZKY, A.: Sensors 101: 3D Sensing. 2021.`https://www.tangramvision.com/blog/sensors-101-3d-sensing`

[2] StereoLabs: ZED 2i Stereo Camera. `https://www.stereolabs.com/en-sk/store/products/zed-2i`

[3] KIRSTEN, E. et al.: 3D Data Acquisition Using Stereo Camera. *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, pp. 9214–9217. 2018.

[4] ZLOCKI, A. et al.: Definition and Application of a Test Methodology for Lidar Sensors. *ATZelectronics worldwide*, vol. 16, pp. 46–49. 2021

[5] Vzense: Vzense DCAM550 ToF Camera User Manual. `https://www.vzense.com/uploads/files/20230112/ce18b93bee68af7a9aa8e4307ba16217.pdf`

[6] WANG, P.: Research on Comparison of Lidar and Camera in Autonomous Driving. *Journal of Physics: Conference Series*, vol. 2093, no. 012032. 2021.

[7] JANG, M. et al.: A Comparison and Evaluation of Stereo Matching on Active Stereo Images, *Sensors*, vol. 22, no. 3332. 2022.

[8] ABDELSALAM, A. et al.: Depth accuracy analysis of the ZED 2i Stereo Camera in an indoor Environment, *Robotics and Autonomous Systems*, vol. 179, no. 104753. 2024.

[9] RUSTLER, L. et al.: Empirical Comparison of Four Stereoscopic Depth Sensing Cameras for Robotics Applications, *IEEE Access*, vol. 13, pp. 67564-67577. 2025.

[10] GARIGIPATI, B. et al.: Evaluation and comparison of eight popular Lidar and Visual SLAM algorithms, *2022 25th International Conference on Information Fusion (FUSION)*, pp. 1-8. 2022.

[11] Python Docs: tkinter – Python interface to Tcl/Tk. `https://docs.python.org/3/library/tkinter.html`

[12] Python Docs: threading – Thread-based parallelism. `https://docs.python.org/3/library/threading.html`

Received March 5, 2025, accepted May 15, 2025

BIOGRAPHIES

Ľubomír Urblík is a PhD student at the Department of Cybernetics and Artificial Intelligence (DCAI), FEEI, TUKE. His research focuses on using containers in edge environments, specifically, AI and data processing.

Jakub Blizman graduated (Bc) in 2024 at the Department of Cybernetics and Artificial Intelligence, FEEI, TUKE.

Erik Kajáti is an assistant professor at the Department of Cybernetics and Artificial Intelligence (DCAI), FEEI, TUKE. His research focuses on Industry 5.0, particularly sustainability, resiliency, and human-centricity, with applications in the energy sector, healthcare, and manufacturing.

Peter Papcun is an associate professor and the head of the Department of Cybernetics and Artificial Intelligence (DCAI), FEEI, TUKE. His research is focused on Industry 4.0, Cyber-Physical Systems (CPS), and Industrial Internet of Things (IIoT).