EVALUATION AND COMPARISON OF DIFFERENT TIME OF FLIGHT CAMERAS FOR OUTDOOR APPLICATIONS

Z. Qiu[1,2],* J. Martínez-Sánchez[1], P. Arias[1]

[1] CINTECX, Universidade de Vigo, Applied Geotechnology Group, Vigo, Spain - (zhouyan.qiu, joaquin.martinez, parias)@uvigo.es
[2] ICT & Innovation Department, Ingeniería Insitu, Vigo, Spain - zhouyan.qiu@ingenieriainsitu.com

KEY WORDS: Time of Flight camera, Mobile mapping, Azure Kinect, SLAM, depth imaging, sensor evaluation

ABSTRACT:

Time-of-Flight (ToF) cameras have gained prominence in robotics, augmented reality, and gesture recognition due to their cost-effective direct measurement of 3D environments. However, their outdoor applications remain limited, mainly due to challenges like sunlight interference. Through systematic testing under challenging outdoor conditions, we aim to assess the suitability of ToF cameras, specifically Azure Kinect and Pmd Monstar, in smart city contexts and contribute to the state-of-the-art and future trends of 3D sensing technology. Our experiments focus on three high-reflectivity cases: license plates, reflective road marking paint on cement and asphalt boards, and traffic cones. Results indicated that Azure Kinect offered a longer measurement range but was more susceptible to flying pixels. Pmd Monstar provided more stable depth measurements and was less sensitive to flying pixels. Differences in performance were attributed to their modulation frequencies and the distinct approaches to handling low-confidence points. By addressing the identified limitations and challenges, researchers and engineers can enhance ToF camera capabilities, ultimately improving their performance and expanding their applicability in outdoor transportation, autonomous driving, and other related smart city fields.

1. INTRODUCTION

Time-of-Flight (ToF) cameras have emerged as a powerful tool in various fields, including robotics (Li and Liu, 2019), augmented reality (Gu et al., 2021), and gesture recognition (Suarez and Murphy, 2012), due to their capacity to measure object reflection times and provide cost-effective, direct measurement of 3D environments. ToF cameras have become increasingly popular because of their unique ability to capture depth information in real-time with relatively low computational requirements (Foix et al., 2011), making them an attractive option for many applications.

Despite their advantages, the use of ToF cameras in outdoor scenarios (Qiu et al., 2022b), such as transportation and autonomous driving (Yurtsever et al., 2020), remains limited, primarily due to challenges posed by sunlight interference, which can adversely affect the accuracy and reliability of the acquired depth data (Kurillo et al., 2022). To address these limitations and unlock the potential of ToF cameras in such contexts, this study aims to assess and compare their performance under various outdoor conditions involving high-reflectivity objects that are crucial for traffic safety and management. The study also explores the performance of ToF cameras when dealing with different types of high-reflectivity objects, such as license plates, reflective road marking paint, and traffic cones.

By examining the performance of ToF cameras in these challenging conditions, the research aims to provide valuable insights into their potential applications in transportation and autonomous driving, as well as inform future developments in the field. The findings from this study are expected to contribute significantly to the ongoing efforts to improve the performance of ToF cameras, ultimately enabling their widespread adoption in smart city applications, such as city infrastructure inventory, city planning, urban design, and autonomous driving, among others.

The remainder of the paper is structured as follows: Chapter 2 provides an overview of the camera and outlines its experimental design. Chapter 3 addresses common challenges faced by ToF cameras and presents our proposed solutions. In Chapter 4, we conduct a comparative performance analysis and evaluation for three distinct cases. A detailed discussion of the experiment results are presented in Chapter 5. Finally, Chapter 6 offers a summary and concluding remarks.

2. SENSOR SPECIFICATIONS AND EXPERIMENTAL DESIGN

2.1 Sensor Specifications

The cameras we are testing in this study include Azure Kinect, and Pmdtec Monstar. In 2019, Microsoft launched Azure Kinect (Microsoft, 2023), which is a multi-sensor system containing RGB camera, ToF camera, gyroscope and accelerometer, and microphone array. Besides, the CamBoard pico Monstar is one of the most powerful and versatile depth sensing devices developed by pmdtechnologies (pmdtechnologies, 2023). Table 1 listed different characteristic of these cameras. In this study, we used Azure Kinect ToF camera's NFOV Binned mode with a 320 x 288 px resolution for all the three cases and used WFOV Binned mode with a 512 x 512 px resolution for Case 3: traffic cone.

2.2 Experimental Design

Azure Kinect and Pmd Monstar were securely attached using nano glue, and affixed to a durable 360° aluminum alloy pan-tilt head, allowing for full rotation of the translation base and vertical arm in both horizontal and vertical directions. The entire setup was then mounted on a tripod. This integration is

---

* Corresponding author

Features	Azure Kinect	Pmd Monstar
RGB Camera	3840 x 2160 px @30 fps	-
Depth Camera	640 x 576 px @30 fps, 0.5m-3.86m 320 x 288 px @30 fps, 0.5m-5.46m 512 x 512 px @30 fps, 0.25-2.88m 1024x1024 px @15 fps, 0.25-2.21m	352 x 287 px @60 fps, 0.5m-6m
Wavelength	860nm	850nm
Power	External PSU	USB
Synchronization	RGB & Depth internal only	External device-to-device
Dimensions	103 x 39 x 126 mm	66 x 62 x 29 mm
Mass	440g	142g

Table 1. Comparison of different ToF cameras

provides in Figure 1(a). Additionally, the artificial evaluation targets are shown in Figure 1(c), except for Case 1, which involves using an actual license plate from a car outdoors (Figure 1(b)).

3. PROBLEMS ASSOCIATED WITH THE USE OF TOF CAMERAS OUTDOORS

3.1 Phase wrapping

The ToF camera operates by detecting the phase shift between the signal emitted and the signal reflected. The function for distance measurement is the arctangent of the phase φ in the detected signal. Due to the period of 2π, the value has ambiguity at phase $\varphi + 2\pi$ for all $n \geq 0$. Consequently, a modulation frequency f_{mod} corresponds to a maximum range d_{max} determined by the equation:

$$d_{max} = \frac{c}{2f_{mod}} \qquad (1)$$

where: d_{max} is the maximum measuring range without ambiguity, c is the speed of light, f_{mod} is the modulation frequency of the emitted signal (He and Chen, 2019). For a position beyond d_{max}, the actual distance might be $d_p + n \cdot d_{max}$. The phase wrapping issue requires an algorithm to determine the unknown n_p, as described in Equation 2, known as phase unwrapping ((Hansard et al., 2013)):

$$\mathbf{X}_p(n_p) = \frac{d_p + n_p d_{\max}}{d_p} \mathbf{X}_p \qquad (2)$$

where the measured distance d_p equals $\|\mathbf{X}_p\|$, $\mathbf{X}_p(n_p)$ is the unwrapped 3D point, n_p is the number of wrappings.

In this study, the three cases illustrated in Figure 1(c) and Figure 1(b) involve known object sizes. By calculating their dimensions on the depth map, we can determine the value of n_p. It is worth noting that, due to depth map resolution limitations, the maximum value of n_p attainable by this algorithm is 2. Therefore, the computation of actual distance can be simplified to Equation 3 (Qiu et al., 2022a):

$$\mathbf{X}_p = \begin{cases} d_p, & \text{if } n = 0 \\ d_p + d_{max} & \text{if } n_p = 1 \\ d_p + 2d_{max} & \text{if } n_p = 2 \end{cases} \qquad (3)$$
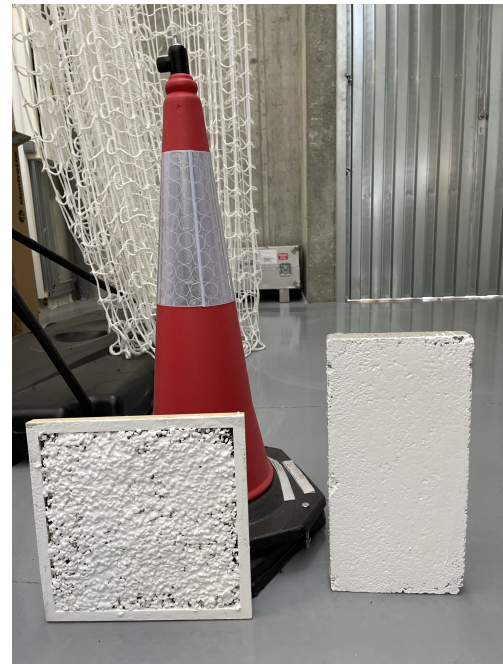
In the subsequent experiments, the outcomes related to the phase wrapping issue have been subjected to a phase unwrapping procedure.

(a) The Sensor Setups

(b) **Case 1: License Plate**, with the specific license plate number obscured

(c) The artificial targets used in the study, including **Case 2: Reflective Road Marking Paint on Cement Board and Asphalt Board and Case 3: Traffic cone**. The cement and asphalt boards are manually crafted and subsequently coated with paint specifically designed for road markings.

Figure 1. Experimental Design

3.2 Flying pixels

In depth images obtained by ToF cameras, flying pixels represent erroneous depth measurements arising from mixed pixels or the overlapping of light from multiple sources. Such inaccuracies transpire when an image sensor pixel captures light from various distances or objects, yielding depth values unrep-

resentative of any genuine surface (Lindner and Kolb, 2006).

We placed targets at varying distances and performed 100 repeated measurements. The measurement outcomes are visualized in Figure 2 and Figure 3. As evident from Frame 0 and Frame 1 examples, numerous flying pixels appear in the image, resulting in a substantial depth standard deviation among the initial 100 images collected and a less accurate depth mean. By continuously measuring three consecutive frames and removing points not present in all three frames, we recalculate the depth standard deviation and depth mean. The figure clearly shows that the processed depth standard deviation is significantly improved compared to the unprocessed data. In this way, we minimize the impact of flying pixels. It is important to note that, as observed in the Figure 3, Pmd Monstar employs a more aggressive strategy to filter out low-confidence points compared to Azure Kinect (Figure 2). This approach results in significantly fewer flying pixels in the depth maps captured by Pmd Monstar. Furthermore, both before and after processing, the standard deviation of the depth maps collected by Pmd Monstar remains substantially lower than that of Azure Kinect. This difference in data handling strategies contributes to the overall performance variations between the two cameras when dealing with highly reflective surfaces and objects with varying surface roughness in outdoor environments.
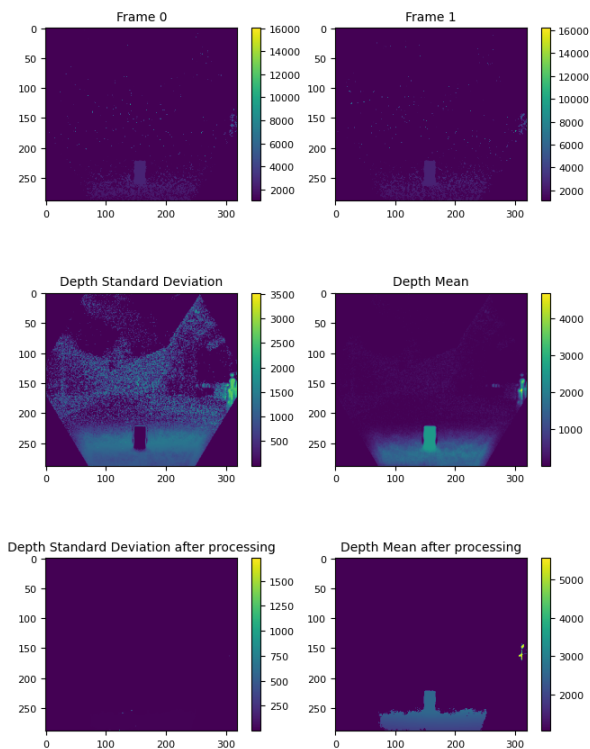

Figure 2. Depth Data Visualization for **Case 3: Traffic Cone** captured by Azure Kinect using 'viridis' colormap, where darker colors signify lower depth values and lighter colors indicate higher values. (Colorbar unit: mm).

4. COMPARATIVE PERFORMANCE ANALYSIS AND EVALUATION OF THREE SPECIFIC CASES

All our experiments were carried out at midday on a sunny day, thoroughly considering the influence of sunlight. In our exper-
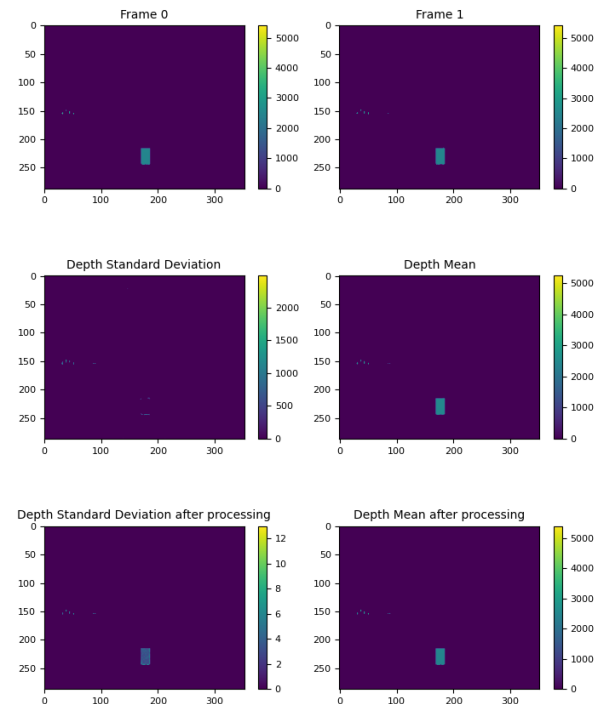

Figure 3. Depth Data Visualization for **Case 3: Traffic Cone** captured by Pmd Monstar using 'viridis' colormap, where darker colors signify lower depth values and lighter colors indicate higher values. (Colorbar unit: mm).

imental investigations, we examine three distinct cases of high reflectivity: license plate, reflective road marking paint on cement and asphalt boards, and traffic cone, which represent three common highly reflective objects in outdoor environments.

4.1 Case 1: License Plate.

Figure 5 presents an example of Azure Kinect and Pmd Monstar cameras capturing a license plate at the distance of around 6 meters. The left image displays the original depth maps, while the right images shows the depth maps processed using three frames. This decision is underpinned by the fact that three frames prove sufficient in eradicating the majority of the flying pixels, while concurrently satisfying the mobility requirements of the vehicles used for outdoor data collection. As evident from Figure 4(a), the depth map obtained by Azure Kinect contains numerous flying pixels. In contrast, the depth map provided by Pmd Monstar (Figure 4(b)) exhibits significantly fewer flying pixels, with minimal differences between the before and after processing results.

Another intriguing finding, as seen in Figure 4(b), is that Pmd Monstar primarily collects data at the license plate's edges. A possible explanation for this phenomenon is that the license plate's reflection under intense sunlight is too strong, thereby interfering with the ToF camera's data acquisition. A similar occurrence is also observed in Case 3: Traffic cone.

In our experiments, we employed Azure Kinect and Pmd Monstar to capture 100 repeated depth images at each measurement point, then obtaining 98 depth maps following the processing outlined in Chapter 3.2. To determine the distance between the license plate and the camera in each image, we calculated the

(a) The depth map captured by Azure Kinect. Left: Frame 1; Right: After removing flying pixels

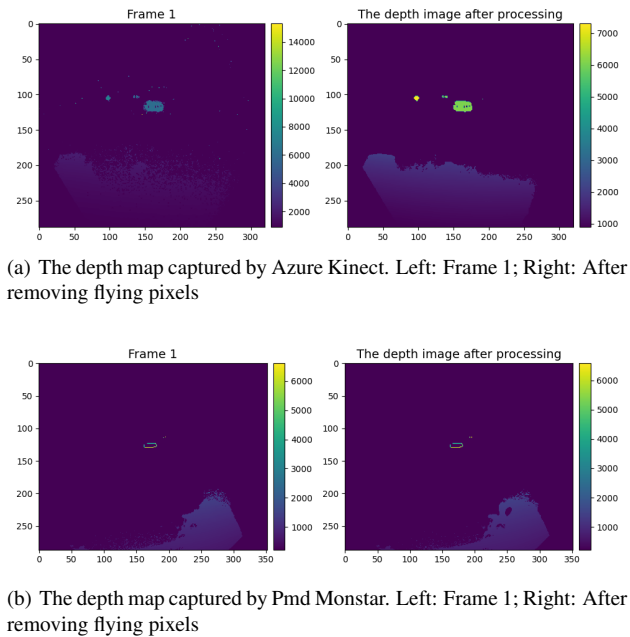(b) The depth map captured by Pmd Monstar. Left: Frame 1; Right: After removing flying pixels

Figure 4. Examples Visualization for **Case 1: License Plate** using 'viridis' colormap, where darker colors signify lower depth values and lighter colors indicate higher values. (Colorbar unit: mm).

average pixel value at the license plate location on each depth images. Subsequently, we computed the standard deviation and mean of the resulting 98 depth values and illustrated them in the Figure 5. As depicted in Figure 5, Pmd Monstar demonstrates a maximum measurement distance of approximately 20 meters, showcasing its capacity for long-range depth sensing. In comparison, Azure Kinect is capable of achieving an even greater limit measurement distance, reaching up to 25 meters. This extended range is complemented by a smaller standard deviation, suggesting a higher level of precision in its measurements. It is important to note that as the distance increases, so does the standard deviation, signifying a relationship between the two variables. This observation highlights the relationship between distance and the precision of the depth measurements in both cameras.
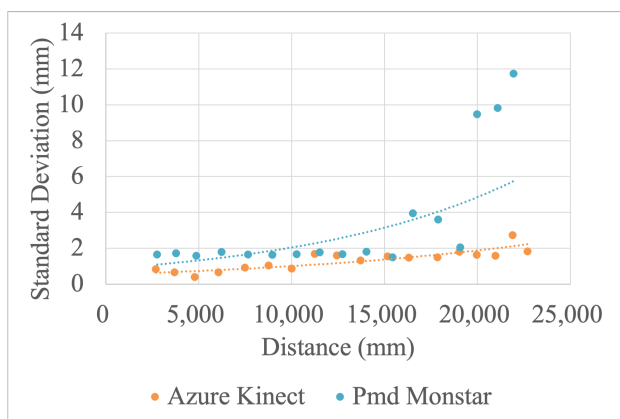

Figure 5. Comparison of Standard Deviation in Distance Measurements for Azure Kinect and Pmd Monstar (**Case 1: License Plate**) with Trend Lines

4.2 Case 2: Reflective Road Marking Paint on Cement Board and Asphalt Board.

In this scenario, we evaluated the performance of employing the same high-reflectivity paint on distinct surfaces. Similar to Case 1, we present an example of Azure Kinect and Pmd Monstar capturing depth images of cement board and asphalt board surfaces. As depicted in Figure 6(d), Pmd Monstar fails to acquire valid data on the asphalt Board surface. This may be attributed to the excessive roughness of the asphalt Board surface, which generates multi-path interference. This results in a low confidence level for the acquired asphalt Board data, prompting Pmd Monstar to filter out these points.

Similarly, we compare the standard deviation of depth images captured by Azure Kinect and Pmd Monstar at various distances. Pmd Monstar is unable to provide depth data on asphalt board, and the effective measurement distance for Azure Kinect is limited to approximately 6 meters. In contrast, when assessing the cement board surface, both cameras exhibit significantly improved performance due to differences in surface roughness. On the cement board surface, Pmd Monstar is capable of measuring distances beyond 4 meters, indicating its ability to handle such surfaces to a certain extent. At the same time, Azure Kinect maintains an impressively low standard deviation of around 2 mm at a measurement distance of nearly 9 meters, demonstrating its precision and reliability on this type of surface.

4.3 Case 3: Traffic cone.

As shown in Figure 1(c), the reflective cone collar is applied to the traffic cones. This collar is manufactured using glass bead material and features a type 3 reflectivity classification: High Intensity Prismatic (Carbide and Engr, n.d.). This level of reflectivity is primarily employed in high-visibility applications, such as traffic signs, construction zone devices, and delineators, ensuring optimal visibility and safety in various lighting conditions.

An interesting aspect of the traffic cone results, as evidenced in Figure 8, is the conspicuous absence of data within the collar of the traffic cone. This can be attributed to the excessively high reflectivity of the collar, so the cameras fail to accurately capture depth information in that particular region. As a consequence, the data points tend to cluster around the periphery of the collar area, highlighting the limitations of the cameras in handling highly reflective surfaces. This issue is also evident in Figure 4(b). However, on the positive side, Figure 9 reveals that the peripheral data acquired is fairly reasonable. By calculating the distance and standard deviation using the obtained data from the marginal portion, we can deduce that the acquired data is both reliable and relatively stable. This suggests that, despite the challenges posed by highly reflective surfaces, depth-sensing cameras can still provide valuable and dependable information from the surrounding areas, which may be useful in various applications and real-world scenarios.

In this case, we also collected the WFOV mode results from Azure Kinect, as demonstrated in Figure 8(b). Figure 9 compares the relationship between distance and standard deviation for depth maps captured by the two modes of Azure Kinect alongside Pmd Monstar.

As observed in the Figure 9, for the traffic cone, the maximum measurement distance in Azure Kinect's WFOV mode is approximately 13 meters, with a corresponding standard deviation
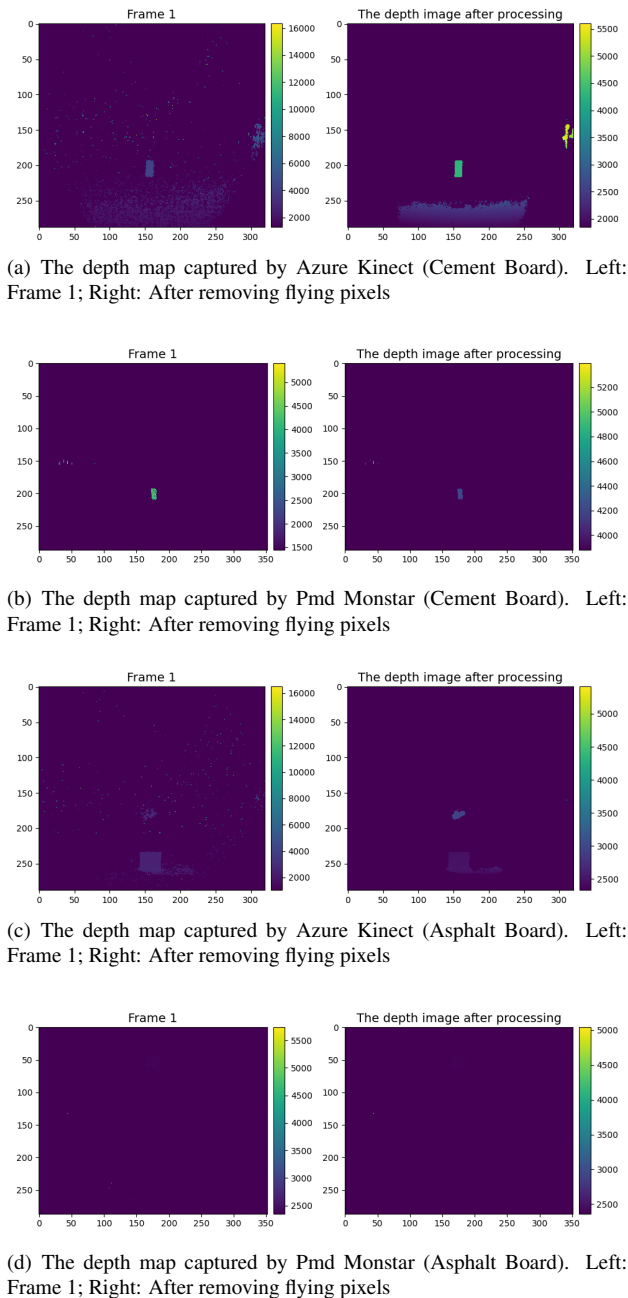
(a) The depth map captured by Azure Kinect (Cement Board). Left: Frame 1; Right: After removing flying pixels

(b) The depth map captured by Pmd Monstar (Cement Board). Left: Frame 1; Right: After removing flying pixels

(c) The depth map captured by Azure Kinect (Asphalt Board). Left: Frame 1; Right: After removing flying pixels

(d) The depth map captured by Pmd Monstar (Asphalt Board). Left: Frame 1; Right: After removing flying pixels

Figure 6. Examples Visualization for **Case 2: Reflective Road Marking Paint on Cement Board and Asphalt Board** using 'viridis' colormap, where darker colors signify lower depth values and lighter colors indicate higher values. (Colorbar unit: mm).

of about 1mm. In contrast, the NFOV mode can achieve a measurement distance of around 20 meters, with a corresponding standard deviation of less than 1mm. Pmd Monstar also reaches a maximum measurement distance of about 20 meters, with a similar standard deviation of around 1mm.

Another intriguing observation is that, as seen in Figure 9, the standard deviation increases with distance for both Azure Kinect modes, whereas the distance measured by Pmd Monstar remains relatively stable, fluctuating between 1mm and 1.2mm. This finding highlights the different performance characteristics of the TOF cameras when operating under varying field-of-view settings and working with highly reflective surfaces.
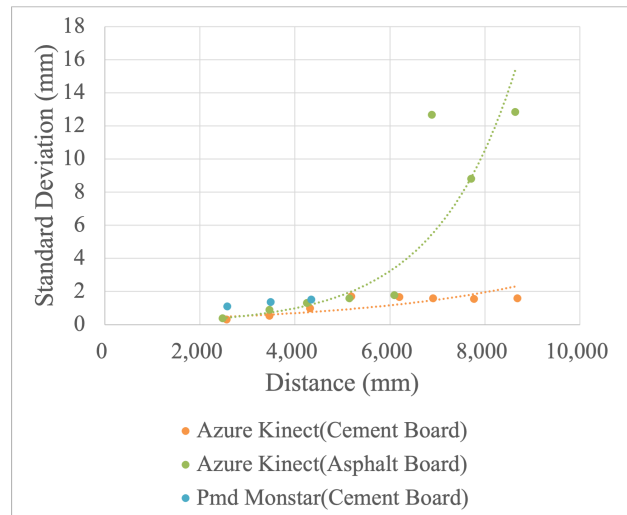

Figure 7. Comparison of Standard Deviation in Distance Measurements for Azure Kinect and Pmd Monstar (**Case 2: Reflective Road Marking Paint on Cement and Asphalt Boards**) with Trend Lines

(a) The depth map captured by Azure Kinect (NFOV mode). Left: Frame 1; Right: After removing flying pixels

(b) The depth map captured by Azure Kinect (WFOV mode). Left: Frame 1; Right: After removing flying pixels

(c) The depth map captured by Pmd Monstar. Left: Frame 1; Right: After removing flying pixels
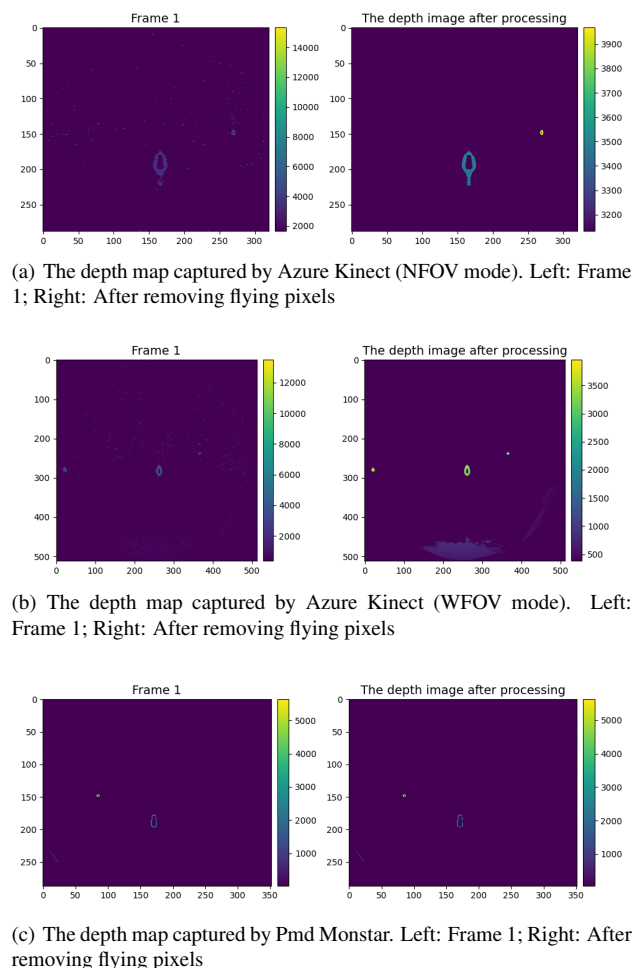
Figure 8. Examples Visualization for **Case 3: Traffic cone** using 'viridis' colormap, where darker colors signify lower depth values and lighter colors indicate higher values. (Colorbar unit: mm).
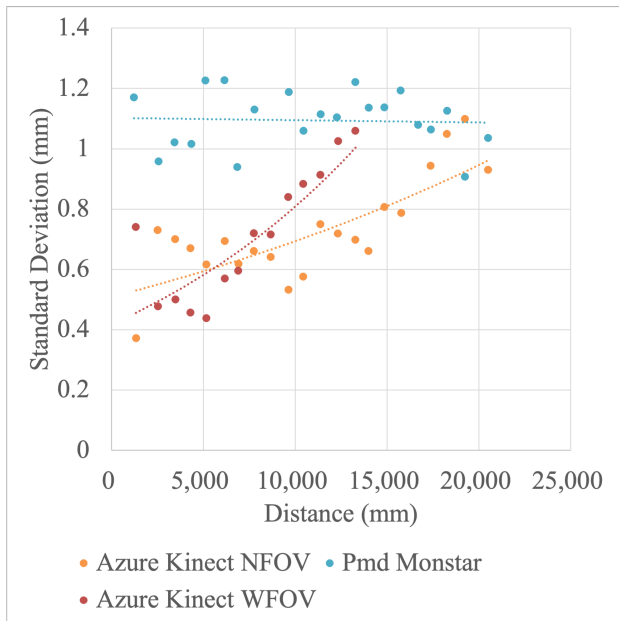
Figure 9. Comparison of Standard Deviation in Distance Measurements for Azure Kinect and Pmd Monstar (**Case 3: Traffic Cone**) with Trend Lines

5. DISCUSSION

In this study, we aimed to investigate the performance of two ToF cameras, Azure Kinect and Pmd Monstar, when encountering highly reflective surfaces in outdoor environments. Three distinct cases were considered: license plates, reflective road marking paint on cement and asphalt boards, and traffic cones with high-intensity prismatic collars. Our experiments validate the reliability of ToF cameras when conducting outdoor measurements on highly reflective objects. The maximum measurement distance is influenced by factors such as the ToF camera's resolution and modulation frequency, as the well roughness and reflectivity of the object being measured.

Our experiments demonstrated that both cameras struggled to provide accurate depth data for the areas with excessive reflectivity, such as license plates and traffic cones. However, the acquired peripheral data was found to be relatively reasonable, reliable, and stable, suggesting that ToF cameras can still offer valuable information in real-world scenarios involving reflective surfaces.

Another notable finding was the impact of surface roughness on camera performance. For instance, Azure Kinect and Pmd Monstar cameras showed improved performance on cement board surfaces compared to asphalt board surfaces, which affected the ToF cameras' ability to accurately capture depth data.

Moreover, we also compared the performance of Azure Kinect and Pmd Monstar. The following observations were made:

1. **Flying pixels**: Both cameras experienced issues with flying pixels, which are erroneous depth measurements resulting from mixed pixels or superposition of light from multiple sources. However, the extent of flying pixels varied between the two cameras. Depth images captured by Azure Kinect exhibited a larger number of flying pixels compared to Pmd Monstar.

2. **Maximum measurement distance**: The maximum measurement distance was influenced by factors such as the ToF camera's resolution and modulation frequency, as well as the surface roughness and reflectivity of the object being measured. Azure Kinect achieved a limit measurement distance of up to 25 meters, while Pmd Monstar reached approximately 20 meters. However, these distances were affected by the specific scenarios and objects under examination.

In conclusion, both Azure Kinect and Pmd Monstar demonstrated strengths and weaknesses when encountering highly reflective surfaces in outdoor environments. While Azure Kinect offered a longer measurement range, it was more susceptible to flying pixels. On the other hand, Pmd Monstar provided more stable depth measurements across varying distances and appeared less sensitive to flying pixels. The observed differences in performance between Azure Kinect and Pmd Monstar cameras may be attributed to their distinct approaches to handling low-confidence points and their varying modulation frequencies. Pmd Monstar appears to adopt a more aggressive strategy when it comes to removing low-confidence points, which might contribute to its reduced susceptibility to flying pixels and its relatively stable depth measurements across varying distances. In contrast, Azure Kinect seems to retain more data points, including those with low confidence, potentially leading to a higher occurrence of flying pixels. Another critical factor influencing the performance of these cameras is their modulation frequency. Differences in modulation frequency can significantly affect the maximum measurement distance, the ability to handle highly reflective surfaces, and sensitivity to surface roughness. The distinct modulation frequencies of Azure Kinect and Pmd Monstar may be responsible for the observed variations in their performance when encountering highly reflective surfaces and objects with differing surface roughness.Further research is needed to determine the optimal camera settings for specific applications involving highly reflective surfaces and other challenging conditions.

6. CONCLUSION

This study systematically assessed the performance of Azure Kinect and Pmd Monstar ToF cameras in outdoor scenarios involving high-reflectivity objects, with a focus on their real-world applicability for transportation and autonomous driving applications. We evaluated the ToF cameras' performance with three high-reflectivity cases: license plates, reflective road marking paint on cement and asphalt boards, and traffic cones.

Our findings demonstrate that ToF cameras offer significant advantages over traditional sensors, such as stereo cameras, by providing direct depth information without the need for complex calculations. In comparison to LIDAR, ToF cameras emerge as a more cost-effective, compact, and energy-efficient alternative, making them an attractive option for a wide range of applications.

Azure Kinect demonstrated a longer measurement range but was more susceptible to flying pixels. Conversely, Pmd Monstar provided more stable depth measurements across different distances and was less sensitive to flying pixels. The observed differences in performance can be attributed to the cameras' distinct approaches to handling low-confidence points and their varying modulation frequencies.

These insights contribute to our understanding of the performance of low-cost ToF cameras in the context of smart cities, particularly when interacting with highly reflective surfaces and objects. The study emphasizes the importance of considering factors such as data processing strategies, modulation frequencies for specific applications within smart city environments. Further research and development efforts are necessary to advance the performance of ToF cameras in these challenging outdoor scenarios, ultimately expanding their utility and effectiveness across a broader range of smart city applications.

ACKNOWLEDGEMENTS

REFERENCES

Carbide, N., Engr, S., n.d. 2014 Traffic Sign Retroreflective Sheeting Identification Guide. *NOTES*, 8, 8.

Foix, S., Alenya, G., Torras, C., 2011. Lock-in time-of-flight (ToF) cameras: A survey. *IEEE Sensors Journal*, 11(9), 1917–1926.

Gu, W., Shah, K., Knopf, J., Navab, N., Unberath, M., 2021. Feasibility of image-based augmented reality guidance of total shoulder arthroplasty using microsoft HoloLens 1. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 9(3), 261–270.

Hansard, M., Lee, S., Choi, O., Horaud, R., 2013. Disambiguation of time-of-flight data. *Time-of-Flight Cameras*, Springer, 29–43.

He, Y., Chen, S., 2019. Recent advances in 3D data acquisition and processing by time-of-flight camera. *IEEE Access*, 7, 12495–12510.

Kurillo, G., Hemingway, E., Cheng, M.-L., Cheng, L., 2022. Evaluating the accuracy of the azure kinect and kinect v2. *Sensors*, 22(7), 2469.

Li, P., Liu, X., 2019. Common sensors in industrial robots: A review. *Journal of Physics: Conference Series*, 1267, IOP Publishing, 012036.

Lindner, M., Kolb, A., 2006. Lateral and depth calibration of pmd-distance sensors. *Advances in Visual Computing: Second International Symposium, ISVC 2006 Lake Tahoe, NV, USA, November 6-8, 2006. Proceedings, Part II 2*, Springer, 524–533.

Microsoft, 2023. Azure kinect dk – develop ai models.

pmdtechnologies, 2023. monstar: picofamily.

Qiu, Z., Martínez-Sánchez, J., Arias, P., 2022a. A novel license plate detection based time-to-collision calculation for forward collision warning using azure kinect. *2022 IEEE 5th International Conference on Image Processing Applications and Systems (IPAS)*, IEEE, 1–6.

Qiu, Z., Martínez-Sánchez, J., Brea, V. M., López, P., Arias, P., 2022b. Low-cost mobile mapping system solution for traffic sign segmentation using Azure Kinect. *International Journal of Applied Earth Observation and Geoinformation*, 112, 102895.

Suarez, J., Murphy, R. R., 2012. Hand gesture recognition with depth images: A review. *2012 IEEE RO-MAN: the 21st IEEE international symposium on robot and human interactive communication*, IEEE, 411–417.

Yurtsever, E., Lambert, J., Carballo, A., Takeda, K., 2020. A survey of autonomous driving: Common practices and emerging technologies. *IEEE access*, 8, 58443–58469.