Towards Robust 3D Perception in Emergency Condition: An Analysis on Commercial Depth Sensors

Geonhwa Son [a], Sangin Lee[a], Taejoo Kim[b], Yukyung Choi [b*]
[a]*Dept. of Artificial Intelligence and Robotics, Sejong Univ., Seoul 05006, Republic of Korea*
[b]*Dept. of Convergence Engineering for Intelligent Drone, Sejong Univ., Seoul 05006, Republic of Korea*
[*]*Corresponding author: ykchoi@rcv.sejong.ac.kr*

***Keywords :** emergency condition, aerosol, perception, depth sensor

1. Introduction

Initial response actions to emergency conditions are crucial for ensuring the safety of nuclear power plants. In response situations such as equipment failure and emergency exit opening, direct operator actions are required. However, working in high-density aerosol environments caused by haze and fog from a nuclear power plant poses risks and hinders task success.

To mitigate this issue, recent researches on nuclear power plant operations [1,2] has adopted unmanned platforms such as robots for worker safety. However, these studies primarily focus on robots for maintenance and inspection and do not address emergency conditions in nuclear power plants. To ensure robust operations in challenging environments, it is important to consider two key factors: understanding the environments that occur in emergency situations, and building a perception system for robots.

In this paper, we first define an aerosol-based emergency level to analyze the robustness of visual perception in challenging environments. Subsequently, we conduct a comprehensive analysis of commercial depth sensors for unmanned platform 3D perception. In addition, we also observed that 3D perception was restricted due to their short wavelength range. Therefore, we investigate the perception possibilities in thermal images, which have relatively long wavelengths, and propose future research based on the consideration of the results.

2. Emergency Definition and Simulation

Nuclear power plants are susceptible to various emergency situations, including high-temperature and high-humidity environments, as well as high radiation exposure, all of which can be caused by nuclear events. Among these challenges, dense aerosol environments significantly reduce visibility, hindering workers immediate response. Unmanned platforms must operate robustly instead of human workers in emergency environments where visibility is limited.

In this study, we define emergency situation as a high-density aerosol environment to simulate visibility restrictions in indoor spaces, while considering real-world disaster scenarios [3,4,5,6,7]. We aim to construct an experimental environment based on the distribution of aerosols in the air. The most significant difference between the defined emergency situation and the normal

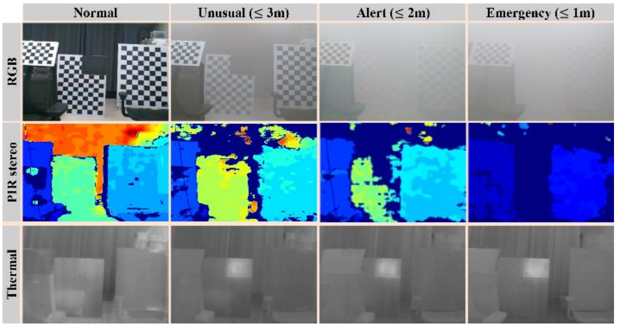

Fig. 1. Comparison of RGB-depth map and thermal image using commercial depth sensor and thermal camera in normal and emergency conditions. Emergency conditions are defined based on the visible distance of checkerboards captured by RGB camera, from left to right: 1m, 2m and 3m.

condition is visibility due to aerosol. Therefore, we define the levels of the emergency situation based on visibility, which changes according to the density of the aerosol. To classify aerosol intensity, checkerboards are installed every 1m within 3m from the camera, as shown in the first row of Fig. 1. Then, we define three levels based on the visibility of each target: usual, alert, and emergency, corresponding to emergency situations relying on aerosol density. Indoor aerosol scenarios are simulated by varying the amount of water vapor injected into the air. For this purpose, a professional haze machine is employed to simulate an aerosol environment by generating a mist of glycerin solution that can be maintained for a relatively long time [8,9].

3. Experiments

In this section, we present a comprehensive analysis of commercial depth sensors in humidity environments. First, we provide an overview of the depth sensors employed in this study. Then, we conduct a comparative analysis of depth map generated by commercial depth sensors. The analysis reveals the limitations of commercial depth sensors. To address the limitation, we conduct thermal camera experiments in high-density aerosol environment. Details about experiments are provided in Sec. 3.2 and Sec. 3.3, respectively.

3.1 Commercial Depth Sensors

Commercial depth sensors utilized in indoor environments are mainly categorized into stereo and

Table I: Commercial depth sensors specification

	Wavelength (nm)	Range (m)	Tech
RealSense D455	865	0.4 – 6	PIR stereo
Azure Kinect	860	0.25 – 5.46	ToF
RealSense L515	850	0.4 - 9	ToF (LiDAR)

time-of-flight (ToF) types [10]. Given the availability of sensors of different size and power consumption, we conduct experiments using three kinds – Passive IR Stereo (Intel Realsense D455), ToF type sensors (Azure Kinect, Realsense L515). Each sensor specification can be shown in Tab. I.

Intel Realsense D455 projects a structured infrared pattern onto the scene and uses stereo matching algorithms to generate depth map. Azure Kinect is time-of-flight camera; it illuminates the scene's objects using an amplitude modulated light source, and the phase delay of modulated signal is measured between emitted and reflected light. Then, phase difference is converted into depth value for each pixel in the image. Intel Realsense L515 is a LiDAR-based depth sensor among ToF type sensor. L515 uses an infrared laser as an active light source. 3D data is produced by calculating the time it takes for the projected signal to bounce off the region in the scene and return to the camera. We evaluate the robustness of these three types of sensors for generalized analysis in the aerosol environments defined in Sec.3.2.

3.2 Depth Sensor Analysis in Aerosol Environments

Comparing across three sensors, depth map consistently shows increased noise with higher emergency levels, as shown in Fig. 2. This finding suggests that the impact of higher emergency levels on

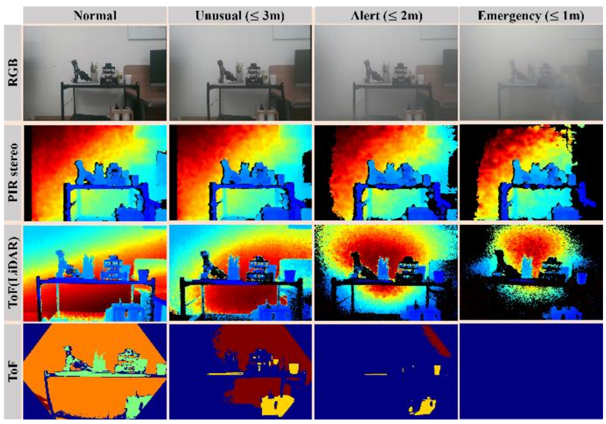

Fig. 2. Comparison of RGB-depth map pairs in normal and emergency conditions. Three-level emergency condition classification based on visible distance from a RGB camera in row 1, with corresponding results from commercial sensors in rows 2-4. From the second row, the depth maps of the Passive IR Stereo (D455), ToF type sensors (L515, Azure Kinect), in that order.

Table II: Performance comparison of commercial depth sensors. The highest performance is underlined

	RMSE (↓)		
	Unusual (≤ 3 m)	Alert (≤ 2 m)	Emergency (≤ 1 m)
D455	<u>29.54</u>	<u>72.10</u>	<u>93.31</u>
L515	83.35	117.82	127.58
Azure Kinect	113.76	123.63	126.23

noise in depth map is not limited to specific sensor type but rather general phenomenon. Especially, beginning at the alert level, L515 has more noise in both background and object region than D455. Similarly, the Azure Kinect shows negligible accurate depth value measurements beginning with the alert level, preventing it from collecting of relevant depth value.

For quantitative comparison of depth maps, we use the RMSE (Root Mean Square Error), commonly used metric to measure depth map quality. As shown in Fig. 3. and Tab. II, we evaluate the RMSE of the depth maps under different aerosol stages. The RMSE is the pixel-wise difference between the ground truth (GT) depth map in the Normal stage and the depth map in each aerosol stage. All commercial depth sensors show gradual increase in error from unusual to emergency level. Specifically, D455 outperforms ToF type sensors, by 34.27 pixel and 22.92 pixel, in the emergency level. Both D455 and ToF type sensors demonstrate similar trend in high-emergency environments, with D455 showing the highest level of robustness, followed by ToF type sensors.

We think the observed results are related to wavelengths used by each commercial depth sensor [11]. To measure depth, ToF type sensors use 850 nm and 860 nm wavelengths, respectively. In comparison, D455 uses infrared projector to emit patterns in 840-860 nm range and stereo NIR up to 865 nm to collect depth information. Therefore, we analyze that D455, which utilizes relatively longer wavelength, shows the highest robustness in emergency environments. In addition, unlike ToF type sensors, the D455 uses stereo depth measurement technique that inherently eliminates information loss during light reflection, ensuring robust operation. Nevertheless, commercial depth sensors commonly suffer from noise in high-density aerosol environments, hindering accurate 3D information acquisition.

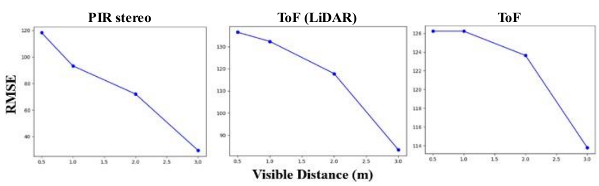

Fig. 3. Qualitative results of commercial depth sensors based on visible distance. We evaluate the performance based on RMSE, where a lower value indicates higher depth quality.

3.3 Robustness Analysis of Aerosols Across Wavelength

In the previous section, we demonstrate the limitations of current commercial depth sensors due to noise of depth map in high-density aerosol environments. We verify wavelength as the reason of this limitation. In addition, we believe that longer wavelength sensors will be more robust in high-density aerosol environments. So we also utilize a thermal camera which has longer wavelength (8-12 um) than commercial depth sensors to validate its robustness.

As we have shown in Fig. 1, thermal images show visibility comparable to normal condition in most regions. These visualization results can be interpreted based on the infrared band wavelengths used by depth sensors and thermal camera. As we can observe, the sensor's robustness to aerosols can be resolved from the long-wavelength, facilitating the construction of cognitive system for unmanned platform in nuclear power plants.

3. Conclusions

In this paper, we analyze commercial depth sensors for robust perception on unmanned platforms in response to the emergent states of nuclear power plants. Specifically, we define a level of emergency to focus on environments where high-density aerosols exist and find that those sensors performances degrade significantly as aerosol density increases. To address this issue, we utilize the thermal sensor and verify that it can secure visibility in an aerosol environment. However, the thermal image sensor is not able to provide depth information, which hinders quantitative comparison. To overcome this limitation, our future work will focus on using longer wavelength sensor for robust perception in challenging environments, especially in high-density aerosol condition.

Acknowledgment

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. RS-2022-00144385)

This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) under the metaverse support program to nurture the best talents (IITP-2023-RS-2023-00254529) grant funded by the Korea government(MSIT)

REFERENCES

[1] E. Zwicker, W. Zesch, R. Moser, A modular inspection robot platform power plant applications, 2010 1st International Conference on Applied Robotics for the Power Industry, p.1-6, 2010.
[2] K. Nagatani, S. Kiribayashi, Y. Okada, K. Yoshida, S. Tadokoro, T. Nishimura, T. Yoshida, E. Koyanagi, M. Fukushima, and S. Kawatsuma, Emergency response to the nuclear accident at the Fukushima daiichi nuclear power plants using mobile rescue robots, Journal of Field Robotics, Vol.30, p.44-63, 2013.
[3] Q. Li, Y. Li, B. Xie, Single image-based scene visibility estimation, IEEE Access, Vol.7, p.24430-24439, 2019.
[4] R. Babari, N. Hautiere, E. Dumont, R.Brmond, N. Paparoditis, A model-driven approach to estimate atmospheric visibility with ordinary cameras, Atmospheric environment, Vol.45, p.5316-5324, 2011.
[5] Z. Lee, S, shang, Visibility: how applicable is the century-old Koschmieder model?, Journal of the Atmosphereic Sciences, Vol.73, p.4573-4581, 2016.
[6] H. Chaabani, F. Kamoun, H. Bargaoui, F. Qutay, A neural network approach to visibility range estimation under foggy weather conditions, Procedia computer science, Vol.113, p.466-471, 2017.
[7] H. Wang, K. Shen, P. Yu, Q. Shi, H. Ko, IEEE, Multimodal deep fusion network for visibility assessment with a small training dataset, IEEE Access, Vol.8, p.217057-217067, 2020
[8] C. Ancuti, C. O. Ancuti, M. Sbert, R. Timofte, Dense-haze: A benchmark for image dehazing with dense-haze and haze-free images, 2019 IEEE International Conference on Image Processing, p.1014-1018, 2019.
[9] C. Ancuti, C. O. Ancuti, R. Timofte, C. De Vleeschouwer, I-HAZE: A dehazing benchmark with real hazy and haze-free indoor images, Advanced Concepts for Intelligent Vision Systems: 19th International Conference, p.620-631, 2018.
[10] H. Sarbolandi, D. Lefloch, A. Kolb, Kinect range sensing: Structured-light versus time-of-flight Kinect, Computer Vision and Image Understanding, Vol.139, p.1-20, 2015.
[11] CP. Colvero, MCR. Cordeiro, JP. Von Der Weid, New proposal for real time measurements of visibility and signal levels in free space optical systems, SBMO/IEEE MTT-S International Conference on Microwave and Optoelectronics, p.344-347, 2005