# INVESTIGATING THE FACTORS THAT INFLUENCE 3D STEREO DEPTH SENSOR NOISE

Terrence Pierce[1,2], Prem Rachakonda[1]
[1]Intelligent Systems Division, National Institute of Standards and Technology, Gaithersburg, MD, USA.
[2]University of Maryland, College Park, MD, USA

## ABSTRACT

3D stereo depth sensors have a variety of applications, including sensing for autonomous vehicles, reverse engineering, and manufacturing automation. The performance of these sensors can be affected by various factors, such as sensor* construction, sensor technology, sensor settings, environment, etc. Understanding parameters that affect sensor output is needed to characterize them and to develop standards.

As machine learning (ML) with 3D point clouds and depth data becomes increasingly prevalent, understanding the data used with these models becomes crucial for improving the adoption rates of such depth sensors. In certain domains, sensor noise and transient effects can become dominant. Reducing noise before using sensor data with ML algorithms is necessary for increased algorithm accuracy.

To characterize depth sensors, we conducted experiments using targets with varying gloss, color, and texture/pattern. Additionally, we studied sensor data quality and noise by exploring sensor parameters such as exposure, gain, and laser power. We found transient effects in both 2D images and depth data captured by sensors. These experiments help inform the operating conditions that could be advised for specific applications and future standards addressing these sensors.

**Keywords:** Stereoscopic Depth, Sensor Noise, 3D Sensing, Standards, Metrology, Machine Learning

## INTRODUCTION

2D and 3D sensors have been employed in various applications, such as autonomous vehicles, autonomous navigation, robotic bin picking, and industrial automation (Yang, et al. 2021[1]), and they are increasingly being aided by image-based machine learning (ML) algorithms. As these application areas and algorithms mature, it is crucial to determine the conditions and parameters that affect the quality of sensor outputs. This work will describe the investigations related to the transient changes and spatial effects of the 3D stereo sensor output that are influenced by sensor settings and target characteristics. This work will enable end-users to make informed decisions on investing in technology that meets the needs of their application.

### State of the art and review of literature

Many research groups have investigated stereo-depth noise, often from the perspective of software/processing-based techniques for noise reduction. Mallick et al. (2014)[2] reviewed noise for a commercial 3D sensor and determined that the broad types of noise often studied are spatial, temporal, and interference. Chatterjee et al. (2015)[3] built a noise model that was compared to experimental data and tested bilateral filtering. Multiple groups have used Markov Random Field models to improve depth data characteristics (Zhang et al., 2015[4]; Zhu et al., 2008[5]). Houshmand et al. (2024)[6] used a neural network based system to reduce active stereo noise.

Some other research groups have studied noise caused by sensors and the environment, more closely resembling this work. Halmetschlager-Funek et al. (2018)[7] compared active stereo technology against other sensors whose performance varies with factors such as different materials and lighting. Pomerleau et al. (2012)[8] discussed the effect of internal temperature (in the context of being affected by the sun) and the reflectance of the target material on characterizing sensor noise. Wang and Shih (2021)[9] used a step method to measure the fluctuations in active stereo measurements at different distances to a flat, planar target. Kamberova and Bajcsy (1998)[10] first modeled the stereo disparity, which was validated experimentally, and conducted target-based tests to determine the scene attributes (such as weak textures) that led to errors

* In this work, a sensor refers to a 3D stereo depth sensor

in depth. Carfagni et al. (2019)[11] analyzed the performance of one 3D sensor and found higher errors associated with laser power set to 0 mW (i.e., with the laser turned off).

**Background and motivation**

There are two primary motivations for performing this work: a) Understanding the effects of sensor noise on 3D data quality (accuracy, precision, and resolution); b) Understanding the sources of sensor noise can help identify ways to improve sensor performance and thereby improve the accuracy of ML algorithms.

This work focuses on stereoscopic depth sensors, which use two cameras and a pattern projector. The work used three commercial active stereo sensors that use different wavelengths of light for illumination, which are listed in Table 1, and their names are intentionally anonymized. The work here informs the possible noise sources to consider when deploying these commercial sensors in real-world applications. These applications could benefit from better noise reduction to obtain more accurate and interpretable data. By understanding the sources of noise in these sensors, users can be informed of situations that cause them and, therefore, devise ways to minimize them. This will also help manufacturers in improving sensor design and development.

Table 1: 3D stereo depth sensors used in this work and their specifications

| Sensor name | Technology | Separate 2D Imager | Frames per second | Light source Wavelength | Depth Image Pixel Resolution* | Depth Working Range |
|---|---|---|---|---|---|---|
| K1 | Active Stereo** | No | 28 to 54 | ≈ 465 nm | 1024 × 768 | 0.23m to 100m |
| R1 | Active Stereo** | Yes (RGB ‡) | 30 to 90 | ≈ 850 nm | 848 × 480 | 0.3m to 3m |
| R2 | Active Stereo** | Yes (RGB) | 30 to 90 | ≈ 850 nm | 848 × 480 | 0.6m to 6m |

* Pixel resolution can be changed, but these were the settings used for experimentation
** Mode can be switched to passive when the emitter/projector is off
‡ Red-Green-Blue

This work is also relevant to improving ML applications, as data quality is crucial for desired ML outputs. For instance, many autonomous vehicles rely on global navigation satellite systems (GNSS) for navigation. In GNSS-denied environments, these vehicles rely on perception sensors such as 3D cameras, depth sensors, and lidars. The output of these sensors is fused and processed using ML algorithms. When an unexpected navigational error occurs, the consequences can be catastrophic. Thus, noise and undesired signals in the sensor output must be mitigated to enable safer operation of such vehicles. By understanding the various noise sources for stereo sensors, as discussed in this study, noise can be reduced before input into an ML model, thus likely improving the model accuracy.

**Parameters that influence sensor data quality**

Several parameters influence the sensor's data quality. In this work, the effects of varying different settings on the sensors were studied. Among several available settings, a subset of common settings across multiple sensors was identified and categorized (see Table 2). The parameters in each row in the table are closely related, though the underlying technology and implementation may differ.

Table 2: List of parameters corresponding to each type of 3D depth sensor.

| Parameter # | Sensor R1 and R2 | Sensor K1 | Description |
|---|---|---|---|
| 1 | emitter_enabled | trigger_1_enabled | Turn the emitter on |
| 2 | exposure | manual_exposure_time | Adjusts the exposure time |
| 3 | gain † | manual_gain | Adjusts the gain |
| 4 | enable_auto_exposure †. | auto_exposure_mode | Toggles the exposure settings from automatic to manual. |
| 5 | laser_power | trigger_1_pulse_width | Adjusts the power of emitter/laser output. |

† These settings cannot be changed through the software development kit (SDK) during data acquisition

- Parameter 1 controls the modes of the sensor's operation – active and passive modes.
    - When the parameter is set to 1, the sensors are said to be in "active" mode, and a light pattern is emitted onto the target.
    - When the parameter is set to 0, the sensors are in passive mode, where no light patterns are generated.
- Parameter 2 controls the exposure time of the sensors in microseconds. The longer the exposure, the slower the capture and the more light enters.
- Parameter 3 controls the amount of gain, which increases the sensor imager's sensitivity.
- Parameter 4 controls whether gain and exposure are determined automatically or manually.
    - In the case of sensor K1, there are 4 settings:
        - 0 corresponds to automatic gain and automatic exposure,
        - 1 corresponds to manual gain and automatic exposure,
        - 2 corresponds to automatic gain and manual exposure, and
        - 3 corresponds to manual gain and manual exposure.
    - In the case of sensors R1 and R2, gain and auto exposure modes were not found to be changeable using the SDK provided by the sensor's original equipment manufacturer (OEM), but these modes can be changed manually using the OEM's own software interface.
- Parameter 5 on all three sensors controls the strength of the emitted signal but is implemented differently.

Parameter 5 for sensors R1 and R2 is in units of power, and it controls the power of the emitter. In contrast, Parameter 5 for sensor K1 is in units of time and controls the duration of each light pulse. The relationship between optical power and pulse width is commonly found using $P_{avg} = P_p \frac{\tau_p}{T}$ where $P_{avg}$ is the average output power, $P_p$ is the peak power, $\tau_p$ is the pulse period, and T is the total period (Burkhart, 2009[12]). The signal becomes visually brighter (in the visible spectrum) as the pulse width increases for sensor K1. For sensors R1 and R2, the infrared (IR) output becomes brighter in the IR spectrum. The increased brightness (in both visible and IR spectra) makes it easier for the sensor to minimize errors in identifying the points on the target.

After studying these parameters, it was observed that the number of unique parameters affecting sensor performance narrows to three: Parameters 2, 3, and 5. Parameters 2 and 3 control the amount of light captured by the sensor, and Parameter 5 controls the amount of light projected onto the target.

**Experimental setup, results, and discussion**

The experimental setup consisted of a sensor with its optical axis oriented approximately normal to the surface of the target (Figure 1). The sensors can obtain data in active and passive modes, i.e., with and without pattern projection. Sensor K1's emitter is a pseudorandom projector in the visible light spectrum, whereas both sensors R1 and R2 emit dots in the IR spectrum (see Figure 2). These patterns assist the sensor in minimizing errors in the correspondence between common points obtained from the two cameras and, therefore, generally, obtains more accurate depth measurements than in passive mode (see Batlle et al., 1998[13]). By varying the strength of each sensor's illumination, correlations to depth noise and measurement accuracy can be studied. These sensors also capture typical 2D images (RGB or grayscale), and thus, the intensity of the images could be investigated as well but were not part of the experiments described here.

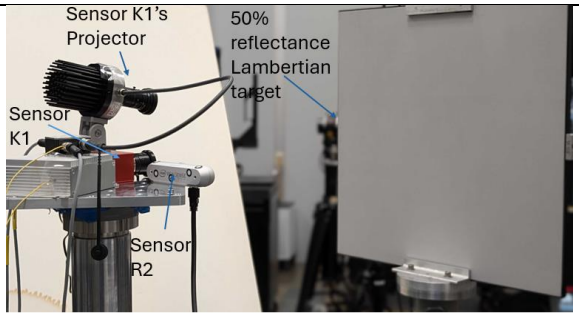
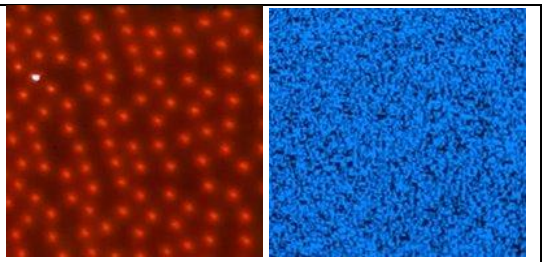



Figure 1. Picture of the experimental setup with sensor K1 on the left and a Lambertian target on the right



a) Sensors R2 b) Sensor K1

Figure 2. Illumination pattern of the depth sensors (R1 has a similar pattern as that of R2)

Data from sensors R1 and R2 were read using the MATLAB[†] SDK from the original equipment manufacturer (OEM). Another method explored to acquire data was to use the OEM's own capture software to record sensor data and to then process the data using MATLAB. However, such data was found to be filtered, which would be undesirable for this study.

Similarly, data from sensor K1 was captured using Python and was analyzed using MATLAB. The Python scripts were called from MATLAB, and once data collection ended, the data was sent directly to the MATLAB script. This method was found to be optimal for this sensor due to repeated, unexpected terminations of MATLAB when reading the sensor's data directly into MATLAB.

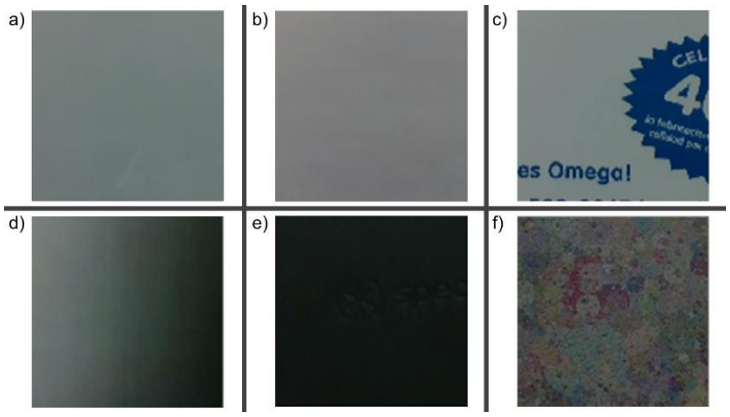


Figure 3. Pictures (200 by 202 pixels) of different targets used in the study. a) white, glossy. b) white, non-glossy. c) glossy and patterned. d) white-to-black gradient non-glossy. e) black, non-glossy, f) spilled coin pattern

Initial experiments involved studying the output of the depth sensors in varying environments. The targets used in these experiments included a glossy white surface, a near-Lambertian white surface, a patterned reflective surface, a white-to-black gradient, a black non-glossy surface, and a spilled-coin pattern (Figure 3). These initial experiments tested the sensor responses to the extremes of glossiness and reflectivity. They also tested for apparent data anomalies in the sensors, e.g., if part of the sensor's field of view did not capture data. However, no significant issues of this nature were found in these experiments.

The subsequent experiments were conducted in a laboratory with a controlled environment. The laboratory had an ambient light level that varied between 330 lx to 370 lx and a temperature that varied between 19 °C to 21 °C. The average light intensity values measured normal to the target varied between 133 lx and 157 lx. A desk fan was used to maintain a consistent temperature in the sensors by treating the large room as a thermal reservoir and cycling in the air at a consistent temperature from the room. Typical temperature readings for sensor K1 and its emitter ranged between 24 °C and 26 °C externally (as measured using a thermometer) and 28 °C to 32 °C internally (as measured using the OEM software). In contrast, sensors R1 and R2 were not found to deviate from the room temperature during these experiments. The laboratory experiments were conducted using a target with a Lambertian gray flat surface with 50% reflectivity (see Figure 1) placed approximately 0.851 m to 0.863 m from the lens of the measuring sensors, which was within the operating range of the sensors. The region of interest was square with a width of 0.260 m to 0.285 m, acquired at the center of the sensors' fields of view to avoid the impact of edge effects.

**Transient changes in sensor RGB and depth output**

We conducted a study to understand if there is any time dependency on the sensor's RGB output. For R1 and R2, using default settings, a time dependency was observed in the RGB intensities upon initializing their pipeline for data collection. The image intensities often settle within 65 to 70 frames or 2.167 seconds to 2.333 seconds. After this stage, the RGB intensities tended to be very stable, with minor variations due to noise. The frame at which stability is reached is indicated by a vertical line near the center of Figure 4 and which denotes the time when all three intensities had reached a steady state. Steady state is determined when the slope of the moving average reaches a threshold value ≈0. This behavior is attributed to the auto-exposure being enabled, as the sensor's exposure needs to adjust to ambient lighting to acquire

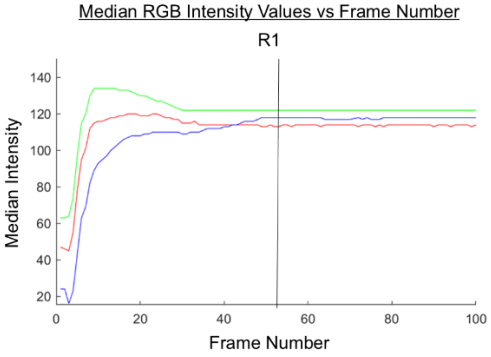


Figure 4. Plots of the RGB intensity values vs the frame number for sensor R1.

[†] Commercial equipment and materials may be identified to specify certain procedures. In no case does such identification imply recommendation or endorsement by the NIST, nor does it imply that the materials or equipment identified are necessarily the best available for the purpose.

images. While these findings are expected, the timing and conditions were not known ahead of time. This behavior can be minimized for applications requiring quick restarts by setting the exposure levels manually based on the previous exposure values.

For depth measurements of R1 and R2, the transient effects are more sudden and binary. In some cases, for highly glossy

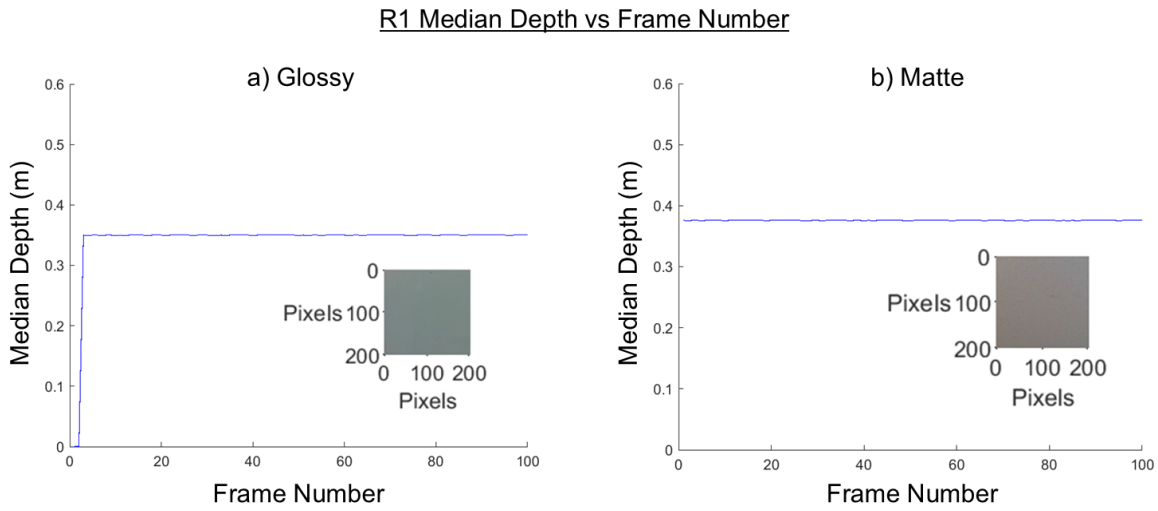



Figure 5. Median depth vs frame number of Sensor R1 aimed at white targets with a) glossy surface, and b) matte surface

surfaces, the sensor takes about 0.1 s (≈three frames) to start receiving data matching the known target distance (see Figure 5a). Essentially, the depth data has a delay, but the duration of this delay is lower than that of RGB images. For matte surfaces, the variation in the depth data is often higher in the first three frames (Figure 5b). After this transient stage, the depth data settles to a steady state with consistent depth measurements. For sensor K1, such transient effects of the same magnitude are not apparent when looking at the corresponding plots.

**Data quality on various regions of the target**

During this study, some systematic variations in data quality (spatial effects) were observed with sensor K1 when acquiring data on a Lambertian target, but not when sensors R1 and R2 were used. Data from sensor K1 had zero depth values on the left side of the target shown in Figure 6.

To investigate this, the OEM of sensor K1 was contacted, and they recommended that the sensor be calibrated. This calibration was performed according to instructions from the OEM, and it resulted in a reprojection error of 0.087 pixels, which is extremely low. Reprojection error is a geometric error corresponding to the image distance between a projected point and a measured one and can be used to ensure correct sensor calibration (Peng et al. 2024, [14]). Despite this calibration, sensor K1 failed to obtain values on the left side of the target, even when cropped to a smaller region of 200 by 202 pixels. This issue is thus far unresolved and needs further investigation.

For sensor K1, it was also observed that a higher pulse width (analogous to a higher brightness or strength of the active mode) resulted in lower noise and a higher data density (percentage of points acquired on the target; see Figure 7a and Figure 7b). Although the data density increased with higher pulse width, the sensor still missed a significant region on the left side of the target.

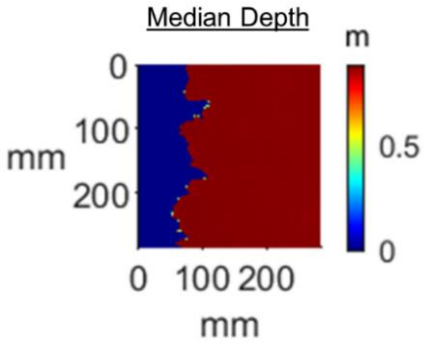



Figure 6. Spatial map of median depth measured with sensor K1. Regions of dark blue correspond to a depth of 0 m, indicating that the sensor is not getting accurate data.

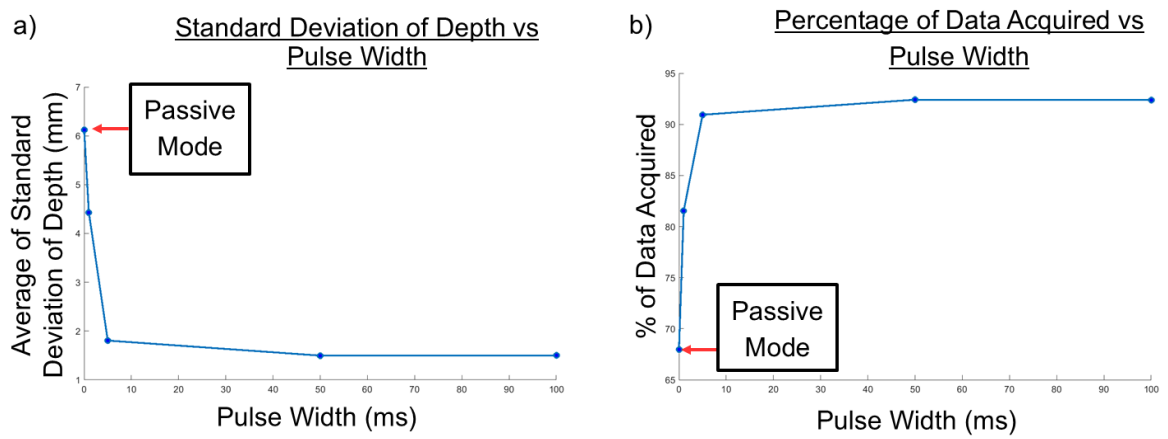



Figure 7. a) standard deviation of depth at different pulse widths for sensor K1. b) percentage of data captured for sensor K1. (Note: A pulse width of 0 ms corresponds to the sensor being in passive mode)

**Effect of gain on sensor noise and data acquired**

The effect of gain on sensor noise and data density was studied by observing data from sensor K1 on the Lambertian target[‡]. This was studied at various exposure levels (including automatic exposure) and pulse widths. However, only the results for automatic exposure using sensor K1 are presented in Figure 8.

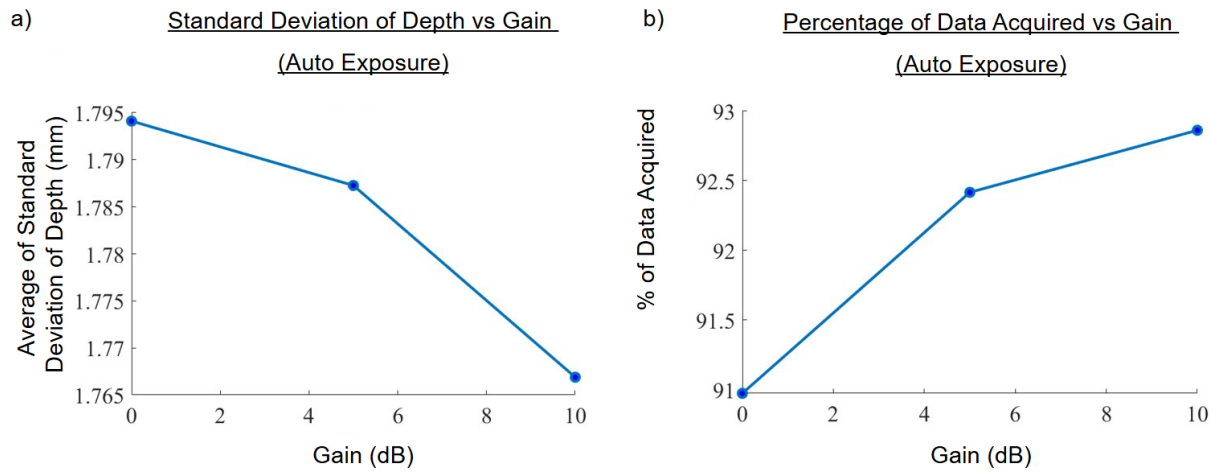



Figure 8. Data parameters of Sensor K1 with auto exposure of 16.136 ms and pulse width set to 5 ms. a) standard deviation of depth vs gain settings. b) percentage of data acquired vs gain settings.

Overall, we observed decreased noise and increased data density when the gain was increased manually. However, this trend reversed when the gain was set to the automatic mode while exposure was also in the automatic mode. For example, in the trials shown in Figure 8, the automatic gain setting indicated a gain value of .00001 dB, resulting in a standard deviation of 1.806 mm and a data density of 90.94% (not plotted in Figure 8). Thus, in some cases, implementing custom optimization algorithms could be the desired approach rather than relying on the sensor's in-built algorithms. While sensor K1 provides settings that might improve these results, not all sensors have these options, or they might optimize different parameters that could be most beneficial for a given application.

[‡] Similar experimentation on R1 and R2 were not conducted since the SDK did not have an option to programmatically vary gain and toggle auto exposure.

## CONCLUSION AND FUTURE WORK

The studies conducted as a part of this work explore sensor noise and data density from commercial 3D depth sensors and offer approaches to improve the data quality. The study on transient effects in these sensors informs the end-user of the duration to wait before a sensor is operational if using an auto-exposure setting. If auto-exposure is not used, the exposure value can be saved and set before sensor initialization to minimize the transient effects (if such an option exists). If the errors in the initial depth frames affect its intended operation, then those frames can be discarded until frame data passes and meets an application-specific threshold.

Turning up the strength of the projector/emitter illumination in the active mode is often beneficial for minimizing measurement variation. Further, target color, gloss, and reflectance can impact sensor noise and data density. Manual implementation of sensor settings might be ideal in specific applications where ambient lighting and targets do not change. Further, the settings to increase the data density on a planar target could be drastically different from those to minimize the standard deviation of the data over the planar surface.

Future work aims to apply optimization techniques to determine the parameters required by an application. This can be compared to the onboard algorithms provided by the sensors and could yield promising results in efficiency and customization. Additionally, studying the effect of sensor data quality on machine learning models can yield insights into methods to improve the accuracy of these models.

## CREDIT AUTHOR STATEMENT

**Terrence Pierce:** Methodology, Formal Analysis, Software, Visualization, Investigation, Writing - Original Draft
**Prem Rachakonda:** Conceptualization, Resources, Supervision, Writing - Review & Editing, Project administration

## ACKNOWLEDGEMENTS

The authors would like to thank members of the NIST Intelligent Systems Division, namely Kamel Saidi, Marek Franaszek, Helen Qiao, and David Schmitt, and Adam Pintar of the NIST Statistical Engineering Division. The authors would also like to thank the Summer Undergraduate Research Fellowship (SURF) and Professional Research Experience Program (PREP) for providing the opportunity to work on this project.

## REFERENCES

[1] Yang, J., Gao, Y., Li, D. and Waslander, S. L., "ROBI: A Multi-View Dataset for Reflective Objects in Robotic Bin-Picking," 2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 9788–9795 (2021). https://doi.org/10.1109/IROS51168.2021.9635871

[2] Mallick, T., Das, P. P. and Majumdar, A. K., "Characterizations of Noise in Kinect Depth Images: A Review," IEEE Sensors Journal 14(6), 1731–1740 (2014). http://dx.doi.org/10.1109/JSEN.2014.2309987

[3] Chatterjee, A. and Govindu, V. M., "Noise in Structured-Light Stereo Depth Cameras: Modeling and its Applications," arXiv:1505.01936 (2015). https://doi.org/10.1007/s11633-018-1114-2

[4] Zhang, S., Wang, C. and Chan, S. C., "A New High Resolution Depth Map Estimation System Using Stereo Vision and Kinect Depth Sensing," J Sign Process Syst 79(1), 19–31 (2015). https://doi.org/10.1007/s11265-013-0821-8

[5] Zhu, J., Wang, L., Yang, R. and Davis, J., "Fusion of time-of-flight depth and stereo for high accuracy depth maps," 2008 IEEE Conference on Computer Vision and Pattern Recognition, 1–8 (2008). http://doi.org/10.1109/CVPR.2008.4587761

[6] Houshmand, P., Staelens, J.-S., Van der Tempel, W. and Verhelst, M., "High-Rate. Compact In-Sensor Denoising for Active Stereo Vision Towards Embedded Depth Sensing," 2024 22nd IEEE Interregional NEWCAS Conference (NEWCAS), 11–15 (2024). https://doi.org/10.1109/NewCAS58973.2024.10666331

[7] Halmetschlager-Funek, G., Suchi, M., Kampel, M. and Vincze, M., "An Empirical Evaluation of Ten Depth Cameras: Bias, Precision, Lateral Noise, Different Lighting Conditions and Materials, and Multiple Sensor Setups in Indoor Environments," IEEE Robotics & Automation Magazine 26(1), 67–77 (2019). https://doi.org/10.1109/MRA.2018.2852795

[8] Pomerleau, F., Breitenmoser, A., Liu, M., Colas, F. and Siegwart, R., “Noise characterization of depth sensors for surface inspections,” 2012 2nd International Conference on Applied Robotics for the Power Industry (CARPI), 16–21 (2012). http://doi.org/10.1109/CARPI.2012.6473358
[9] Wang, T.-M. and Shih, Z.-C., “Measurement and Analysis of Depth Resolution Using Active Stereo Cameras,” IEEE Sensors Journal 21(7), 9218–9230 (2021). https://doi.org/10.1109/JSEN.2021.3054820
[10] Kamberova G., Bajcsy R. “Sensor Errors and the Uncertainties in Stereo Reconstruction”, (1998) https://api.semanticscholar.org/CorpusID:15536700
[11] Carfagni, M., Furferi, R., Governi, L., Santarelli, C., Servi, M., Uccheddu, F. and Volpe, Y., “Metrological and Critical Characterization of the Intel D415 Stereo Depth Camera,” 3, Sensors 19(3), 489 (2019). https://doi.org/10.3390/s19030489
[12] Burkhart, C. “Pulsed Power Engineering Introduction”, Fermilab, Menlo Park, (2009), https://uspas.fnal.gov/materials/09VU/PPE_Introduction.pdf .[Accessed on 2/21/2025]
[13] Batlle, J., Mouaddib, E. and Salvi, J., “Recent progress in coded structured light as a technique to solve the correspondence problem: a survey,” Pattern Recognition 31(7), 963–982 (1998). https://doi.org/10.1016/S0031-3203(97)00074-5
[14] Peng, G., Ren, Z., Gao, Q. and Fan, Z., “Reprojection Error Analysis and Algorithm Optimization of Hand–Eye Calibration for Manipulator System,” 1, Sensors 24(1), 113 (2024). https://doi.org/10.3390/s24010113