2023 IEEE International Conference on Systems, Man, and Cybernetics (SMC)
October 1-4, 2023, Oahu, Hawaii, USA

Evaluation of On-Robot Depth Sensors for Industrial Robotics

1st Odysseus Alexander Adamides
Electrical Engineering Department
Rochester Institute of Technology
Rochester, USA
oaa8092@g.rit.edu

2nd Alexander Avery
Computer Engineering Department
Rochester Institute of Technology
Rochester, USA
aja9675@g.rit.edu

3rd Karthik Subramanian
Electrical Engineering Department
Rochester Institute of Technology
Rochester, USA
kxs8997@g.rit.edu

4th Ferat Sahin
Electrical Engineering Department
Rochester Institute of Technology
Rochester, USA
feseee@rit.edu

2023 IEEE International Conference on Systems, Man, and Cybernetics (SMC) | 979-8-3503-3702-0/23/$31.00 ©2023 IEEE | DOI: 10.1109/SMC53992.2023.10393956

***Abstract*—This work evaluates a Continuous Wave (CW) Time-of-Flight (ToF) camera, Stereoscopic camera, and LiDAR to determine if they are potential candidates for point-rich on-robot sensing in Speed and separation monitoring (SSM) applications. These experiments characterize the static and dynamic behaviors of the sensors while mounted on-robot. From these tests, it was found that ToF and Stereo cameras exhibit better performance to their more expensive LiDAR counterpart. Specifically, it was observed that the ToF camera demonstrated better depth accuracy while the Stereo camera generated better 3D reconstruction accuracy. Overall, ToF and Stereo Cameras demonstrate that with continued innovation and integration, these sensors could become the building blocks to point rich on-robot SSM.**

***Index Terms*—Time-of-Flight, Stereoscopic, Motion Tracking, Angle Dependent Reflectivity, Radius Reconstruction, Speed and Separation Monitoring, Sensor Evaluation.**

I. Introduction

Robotics has increased productivity in manufacturing for decades. The tasks these robots initially performed were repeatable and straightforward. However, as robotic innovations in control and perception have increased, their tasks and proximity to humans have also increased. Now, many robotics companies are manufacturing robots that meet the standards the International Standards Organization (ISO) set to officially designate their robots as collaborative. The ISO/TS 15066:2016 standard outlines four collaborative approaches: safety-rated monitored stop, hand guiding, speed and separation monitoring, and Power and Force Limiting (PFL) [11]. Of these four approaches, companies like Kuka, Universal Robotics, and ABB use PFL. This ISO standard approach requires the companies to build torque and current sensing into their robots to detect when the manipulator robot has unintentionally struck an obstacle or human. The impact force limits are illustrated in [11], but the robot settings must be adjusted for each specific use case [21]. In the event of a collision, the robot moves into a safety-rated monitored stop. In this state, the robot is still powered on but disabled. This is different from an emergency stop because it does not require a complete robot restart to escape this state. This standard approach has been developed and integrated into many industrial-grade robots like the UR-10 [24] and LBR iiwa 7 R800 [12]. Though PFL is currently integrated into many off-the-shelf industrial robots, there has also been extensive research into Speed and separation monitoring (SSM). [1], [3], [9], [13], [14], [22]. This ISO rated approach requires the robot to know the distance and velocity between the robot and the human in the collaborative workspace at all times. If the separation distance becomes too small or the velocity becomes too fast, the robot enters a safety-rated monitor stop. The distinct differences between PFL and SSM are the implementation requirements and the effects on human interaction. In SSM, the robot or the workspace must be outfitted with a perception system to monitor the speed and separation between the robot and the human. In contrast, the PFL cobots must monitor their joint states for torque and current spikes. The key benefit of SSM is that a well-calibrated and integrated perception system significantly reduces human impacts from occurring [13]. PFL is a reactive collaboration approach concerning collision, whereas SSM is a preventative one [14]. This paper will focus on the characterization of point rich 3-D perception sensors to determine if their performance is a good fit for SSM. The remainder of the paper consist of the following sections: Literature Survey, Selected Sensors, Design of Experiments, Results and Discussion, Conclusion, and Future Work.

This material is based upon work supported by the National Science Foundation under Award No. DGE-2125362. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

II. Literature Survey

Within the field of SSM research, there are extensive evaluations of sensing methodologies such as ultrasonic, Infrared, Stereoscopic camera, Time-of-Flight (ToF), LiDAR, and Radar [4], [6], [23]. Additionally, the research generally focuses on how the sensors are integrated into the collaborative workspace. Sensor integration approaches fall into on-robot and off-robot sensing [14]. Off-robot sensing focuses on placing the perception system around the external environment, looking into the collaborative workspace. On-robot sensing utilizes sensors mounted directly to the robot looking out at the collaborative workspace. In [14], off-robot sensing uses point-rich sensors like 3-D ToF Cameras, Stereo Cameras, and LiDARs to capture extensive data on the entire collaborative workspace. Conversely, on-robot sensing is beneficial in highly occluded settings. It is generally performed using lower resolution perception methods like ultrasonic, infra-red, and 1-D ToF Sensors [8], [15]. These distinctions for on-robot sensing were based on the limitations of data-rich sensors, as they were too bulky, consumed too much energy, and required too much computational power. However, in recent years, sensor companies like Intel and Microsoft have begun to develop 3-D perception sensors with lower power and smaller form factors for integration into industrial and automation applications. Furthermore, the rise of embedded Linux Processors like the 'NVIDIA Jetson Platform,' 'Raspberry PI,' and NXP 'i.MX 8' have brought higher processing capabilities to the application endpoint. The overlap between computation and perception has made it feasible to mount Stereo Cameras, LiDARs, and ToF Cameras on robots to achieve point-rich on-robot sensing. These advancements create potential for on-robot sensing with high-resolution returns from the environment. Additionally, these sensors have more expansive field of view (FOV) capabilities, increasing their perception coverage. Increased coverage could lower the chances of occlusions and collisions in the collaborative workspace. Lowering the probability of collision incidences increases the quality of life of the workers who interact with the robots in these work-spaces. As industry and researchers begin to integrate these perception sensors on robots, it is essential to understand the trade-offs between different perception modalities in the SSM application. The robot path, the sensor angle on the robot, and even the type of materials found in the environment can all impact the behavior of these sensors [7]. This paper focuses on characterizing the performance of the Intel Realsense D435i stereoscopic camera [10], Azure Kinect DK 3-D ToF Depth Camera [17], and the Ouster OS0-32 LiDAR [20] when mounted on the tool center point (TCP) of a moving robot. The characterization of these sensors will help determine the considerations necessary for building the next generation of on-robot SSM perception systems. It is crucial to align the performance of the chosen sensor node to the particular SSM application to increase the safety of workers in the collaborative workspace.

III. Selected Sensors

A. Stereoscopic Camera

The D435i is one of the many stereo cameras in the Intel Realsense family. This version contains one 2MP RGB camera with two depth cameras that generate an 87° horizontal by 58° vertical FOV [10]. The D435i has a maximum output resolution of 1280x720, a maximum frame rate of 90 FPS, a clamined sub 2% accuracy error at 2 meters, and a minimum target distance of 280mm at maximum resolution. The operating range for this sensor is 0.3 to 3 meters. These characteristics make the D435i suitable for SSM, as the sensor can accommodate a 3-meter workspace with adequate spatial resolution at a quick refresh rate. It is important to note that trade-offs were made in testing to optimize these characteristics for the SSM experiments performed in this paper, as shown in I. For instance, it was found that best balance between resolution and frame rate in the experiments was 640x480 at 60 FPS.

B. Continuous Wave Time-of-Flight Camera

The Azure Kinect DK (AKDK) is an evaluation platform that combines a 12 MP RGB Camera and 1 MP CW ToF Depth camera into a single platform. This sensor has a 120x120 FOV, 1024x1024 resolution, and a systematic error of less than 11mm plus 0.1% of distance without multipath [17]. The published operating range for this sensor is 0.5 to 3.86 meters in unbinned mode and 0.25 to 2.88mm in binned mode. In order to maintain real-time operating performance, the AKDK was tested at 15 FPS in 512x512 binned mode for the experiments in this paper. 30 FPS and higher resolutions were found to inconsistently drop frames which is not acceptable in a safety critical application.

C. 360° Spinning LiDAR

The Ouster OS0-32 LiDAR is a production-grade near-range LiDAR with a 90° vertical field of view, 32 channels of vertical resolution, and 2048 pixels of horizontal resolution. It only weighs 500g and has a 0.25 to 35m sensing range. The OS0 states a +/- 10mm precision with a 10Hz or 20Hz rotation rate [20]. For the experiments in this paper, the OS0 was run at 20 Hz rotation, 32 vertical channels, and 1024 horizontal pixels, as seen in Table I.

Fig. 1: The D435i, AKDK, and OS0 respectively in their mounts.

D. Motion Capture Ground Truth

The ground truth used for the experiments performed in this work was gathered from an OptiTrack motion capture system. The motion capture system uses twelve Flex 13 OptiTrack cameras to generate a workspace that tracks retro-reflective markers on rigid bodies in the workspace [18]. The system can track these targets with sub-millimeter error at 120 FPS. Additionally, the number of cameras in the system limits the opportunity for occlusions in the workspace. The system only needs 3 of the 12 cameras to locate the position and orientation of any rigid body target in the workspace. The motion capture measured translation and orientations of each rigid body. These poses were used to define the error measured by each experimental depth sensor.

TABLE I: Sensor Parameters as used in Experiments

Sensor Parameters	**Sensor Types**		
	Intel D435i	***Microsoft AKDK***	***Ouster OS0***
Sensing Modality	Stereo	ToF	LiDAR
Resolution	640x480	512x512	1024x32
FOV (°)	87x58	120x120	360x90
FPS	60	15	20
Range (m)	0.3 - 3.0	0.25 - 2.88	0.25 - 35
Spec'd Error	$<$2% at 2m	$<$11mm + 0.1%	5cm
Weight (g)	75	440	500
Size (mm)	90x25x25	103x39x126	87 d x 74.2 h
Cost ($)	334.00	359.10	6,000.00

IV. Design of Experiments

The experiments in this paper cover distances and motions commonly performed by manipulator robots in a collaborative workspace. Static and dynamic motion tests were performed on all three experimental depth sensors with stationary targets and then compared with the ground truth target locations provided by the OptiTrack system. A Universal Robotics UR-10 Robot was used as the on-robot sensing platform to manipulate the sensors through the different paths and positions in testing.

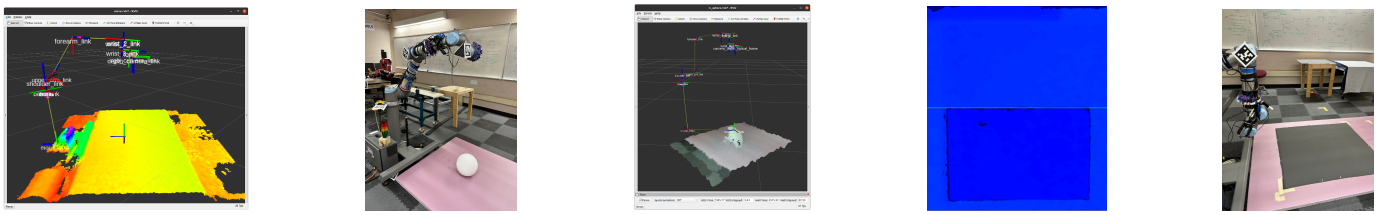
A. Calibration

Before experimentation, a 'eye-on-base' calibration was performed using the 'easy_handeye' calibration package [16]. This calibration involved placing a rigid body target on the end effector and capturing a variety poses to create a frame transformation from the motion capture frame to the robot frame. This process was executed once before any experimental sensor calibration was done. Once the robot motion capture calibration was performed, the experimental depth sensor was mounted to the end effector for extrinsic calibration. Extrinsic calibration for the D435i and AKDK used the 'easy_handeye' calibration package [16] to generate the frame transformation matrix between the camera and robot frames. The process used the RGB camera in the sensor module to take pictures of a 2D checkerboard target on the floor. The robot was moved to different positions and orientations for each capture. The pictures of the checkerboard and the poses of the robot in the different configurations were fed into the 'easy_handeye' calibration package to then generate the camera to robot frame transformation. With this transformation, the transform tree could be constructed from RGB camera frame to depth camera frame, through all the robot joints, to the motion capture system. Full system calibration was required to prevent the error between experimental depth camera measurements and the ground truth location of targets determined by the motion capture setup. Since the LiDAR was lacking an RGB camera, a separate method was used to determine its pose within the workspace. The laser cut LiDAR mount was outfitted with retro-reflective dots to identify the LiDAR as a rigid body in the OptiTrack motion capture software (Motive) [19]. Once the rigid body was created, the center of the rigid body, also known as the pivot point, was adjusted in Motive to match up with the optical center of the LiDAR. Once the rigid body was calibrated to the LiDAR, no eye-in-hand calibration was needed as the LiDAR rigid body position and orientation could be used directly from the motion capture system.

B. Flat Board Tests

First, the sensors went through flat board characterization tests to determine their performance for different target distances and velocities. The static flat test consisted of the UR-10 mounted sensor moving in 100mm increments from 1100mm above a foam board on the ground down to 200mm above the foam board. The foam board was chosen due to its smooth and even surface with a medium level of reflectivity. The ground truth location of the foam board was measured by placing three retro-reflective markers on the board and defining them as a rigid body in Motive. This position and orientation were saved as a static transformation and published into ROS along with the existing motion capture origin, robot, and camera frames as illustrated in Fig. 2a. Each static shot was an average of the captured frames from the sensor for one second. This approach allowed the FPS of each sensor to make an impact on the measurement at each static location. The D435i benefited most from this parameter, as seen in Table I, due to its 60 Hz frame rate. The robot rested at each position for 4 seconds to ensure that the arm was stable before taking the captures. The data collected in the static flat tests were 1-D distances based on the average of a centered region of interest (RoI) on each sensor. This RoI was 5x5 pixels for the Realsense and AKDK and 5x1 pixels for the OS0 LiDAR. The data extracted from these captures were accuracy error μ, measured concerning ground truth, and the standard deviation σ of the measurements across the 1-second recording at each static position.

Next, the sensor was tested across the same distance range with a linear dynamic motion. This test consisted of moving from 1100mm to 300mm towards and away from the foam board. These movements were done at constant velocities of 250mm/s and 500mm/s. The data captured in this experiment were collected at the 700mm point in the motion to guarantee the robot was traveling at a constant velocity. The measured depth, mean accuracy error μ , and standard deviation σ, were captured at this point for +250mm/s, -250mm/s, +500mm/s, and -500mm/s.



(a) (b) (c) (d) (e)

Fig. 2: (2a) AKDK static flat floor recording example. (2b) Experiment setup for static and dynamic data recordings. (2c) Capture from static sphere recording setup. (2d) Capture from static edge precision setup. (2e) Robot setup for ADR testing.

C. Static Sphere Radius Reconstruction Accuracy

As seen in Fig. 2b, the static and dynamic tests were then conducted again but with a 243mm diameter sphere instead of a flat foam board. This target was chosen because it closely resembles a human head, which is a common obstacle in collaborative work-spaces. For these sphere tests, an OptiTrack rigid body target was rested on top of the sphere to determine its XYZ location for the motion capture system. The sphere was then aligned directly underneath the experimental camera on the robot TCP such that the axis of motion was normal to the floor and centered at the top of the sphere. Once this transform was captured, the marker was removed. Fig. 2b is an example capture taken in RViz during D435i testing. The point clouds collected during the sphere tests were used to perform 3-D reconstruction characterization of the experimental sensors. The 3D metrics used in this paper were based on the radius reconstruction accuracy (RRA) used in [7]. RRA compares the point cloud constructed sphere radius with a ground truth radius based on the known dimensions of the target. The radius measurement is obtained from the point cloud by averaging the Euclidean distance of each point registered on the surface of the sphere to the center of the sphere, which is known via the motion capture system. Points with a depth of 0 or NaN were ignored, and captures were only accepted if $>$50% of all possible points were registered on the surface of the sphere. For both the static and dynamic sphere tests, data was collected over 1 second and averaged. The reported metrics for the experiment were average error μ, and standard deviation σ, of these measurements.

D. Edge Precision

Following the sphere tests, edge precision tests were performed following an extension of the procedure outlined in [7]. The robot path setup for this test was similar to the static test illustrated previously. In this case, the target was the edge of a matte-gray corrugated plastic board. The board edge was aligned to the center of the sensor FOV and normal to the robot path. Edge alignment was done to maintain the board edge in the center of the shot at each measured distance. This experiment averaged captures over one second at a static position. The variation in this test compared with [7] was that captures were taken at multiple depths in 100mm steps from 1100mm to 300mm. An ideal line was drawn through the field of view at each captured frame as previously done in [7]. The measured board edge was then compared to the ideal line per pixel separation from that line as seen in Fig. 2d. These pixel offsets were collected and binned into a histogram profile for each 100mm distance. These histograms represented the pixel-wise error variance to the edge of the target. This edge performance characteristic is essential to understand for SSM applications because edge performance indicates the noisiness of the edges in a scene. Noise at transitions, edges, and objects can impact the ability to perform segmentation on that data [5]. The edge precision test was performed dynamically as well. The data capture was done at the 700mm point to ensure constant velocity was achieved. These tests were done at +250mm/s, -250mm/s, +500mm/s, and -500mm/s.

E. Angle Dependent Reflectivity

Following edge precision, the sensors were put through Angle Dependent Reflectivity (ADR) tests as demonstrated in Fig. 2e. These tests were based on the procedure outlined in [7] but on a variation of materials. The depth sensors captured a flat target from 0° to 45° in 5° increments across different material types. Zero degrees in this test was defined as when the sensor was normal to the target. In this SSM-focused application the material types chosen were based on materials that humans may wear in the workspace. These materials included black and gray suit material, fleece, denim, black polyester, flannel, and the foam board as a control material. It was important to test all of these materials, as the perception system chosen for an SSM application should maintain performance regardless of human attire. For this paper the experiments were conducted across the listed angles and materials at a radial distance of 700mm. At each angle, a plane was fit to a cropped region of the point cloud. A plane fit was necessary to isolate performance due to reflectivity from systematic depth error. The distance to the plane was then calculated for each point in the point cloud. The number of points that fall within the σ of the specific sensor to the plane were considered valid. ADR is the percentage of these valid points out of the total number of possible points in the cropped region, based on the sensor resolution. A crop region matching 1/3 by 1/3 of the D534i FOV, or 29x19 degrees, was used across all three sensors to ensure measured points landed within the boundaries of the fabric targets. Additionally, the σ threshold used for each sensor was determined by averaging

(a) (b) (c) (d)

Fig. 3: (3a) Linear path towards the target. (3b) Linear path across the target. (3c) Radial path across the target. (3d) Radial path toward the target.

the performance across all fabrics at 0°. Although, due to the extremely high precision of ToF, its σ was rounded up to 1mm to account for slight wrinkles in the fabric.

F. Dynamic Path Tests

The final sensor tests performed were mock collaborative motions illustrated in Fig. 3. In order to maintain the sensor as the dependent variable in these experiments, the targets used were the same spheres used in the static and dynamic sphere tests. Target placement was designed such that one sphere appeared in the visible path for each motion. The four paths characterized in this experiment were: the linear movement at a target at 250mm/s, linear movement across a target at 250mm/s, radial movement across a target at 60°/s, and radial movement at a target at 60°/s. These were four general movements that could create a human-robot collision. Each target location was designed to stress the limits of the depth sensor to exercise how quickly and accurately the sphere appeared in the frame during each motion. As in the prior tests, depth μ and σ were the key 1-D metrics tracked along the four paths. RRA μ and σ were the 3-D metrics tracked along the paths.

V. Results and Discussion

A. Flat Board Tests

Flat static depth tests revealed that the AKDK generated the best performance with regards to σ and μ across a majority of the testing range as seen in Fig. 4. What was common across all three sensors was that σ decreased as the target distance decreased. The dynamic flat tests in Fig. 5 illustrate a clear impact on the σ of all three sensors. The σ for each sensor in the 500mm/s tests was double that of the 250mm/s tests. There was no clear highest performer from a σ standpoint in the dynamic tests. The boxes in the box style plots consist of a μ center-line and height of +/-3σ. The x-axis represents bins of the tested distances.

B. Static Sphere Radius Reconstruction Accuracy

The stereo camera exhibited the best performance in this testing section. Even though the D435i was not the front runner in the static and dynamic flat depth tests, the D435i generated visibly cleaner point clouds during sphere testing. The ToF camera and LiDAR exhibited noticeable deformation in their point clouds. This visual observation is empirically reflected in Fig. 6.

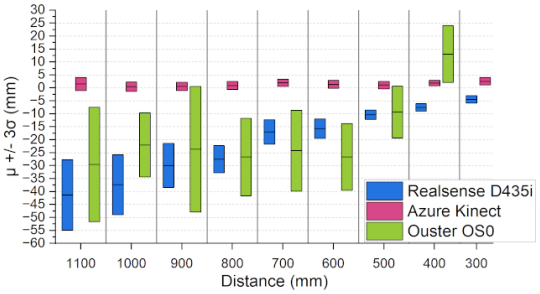


Fig. 4: Static flat depth results from 1100mm to 300mm.

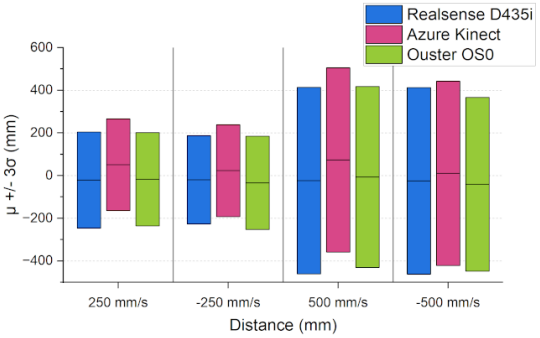


Fig. 5: Dynamic flat depth results at +/-250mm/s and +/-500mm/s.

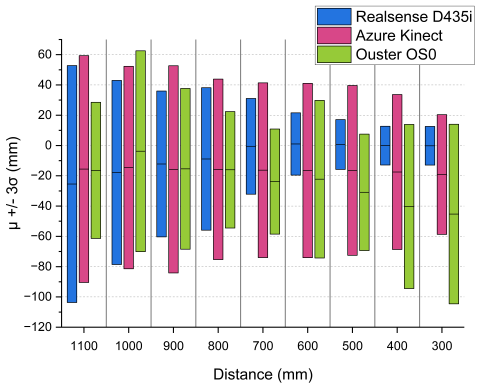


Fig. 6: Static sphere RRA results from 1100mm to 300mm.

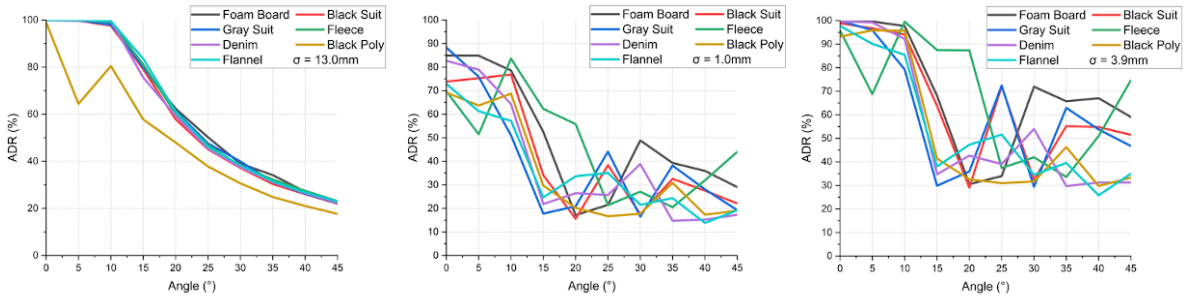


Fig. 7: ADR Results for the D435i, AKDK, and OS0 sensors respectively. The σ threshold is listed in each plot legend

C. Edge Precision

The AKDK and LiDAR performed equally well in this testing section, as seen in Fig. 8. Both of these sensors had much tighter distributions than the D435i. Though the AKDK distribution was tighter than the LiDAR spread, the LiDAR had more zero error points than the AKDK distribution. The dynamic edge results demonstrated similar behavior to dynamic flat depth for the ToF and Stereo Cameras. The 250mm/s distributions contained more zero-pixel error to edge points than the 500mm/s tests. The LiDAR had almost the same number of zero-pixel error points. However, it is essential to note that the LiDAR point resolution along the edge was significantly less than the point resolution of the other sensors. The lack of resolution may explain why the velocity does not impact the LiDAR data as much in this test. The order of performance according to normalized sums for each sensor in the dynamic edge test was the AKDK, the LiDAR, and the D435i per Fig. 9.

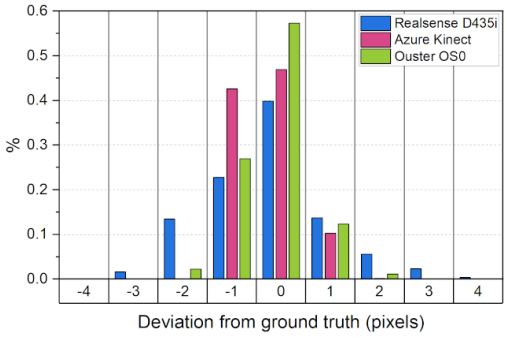


Fig. 8: Pixel spread for each sensor in static edge precision testing

D. Angle Dependent Reflectivity

From a sensitivity standpoint, these tests demonstrated that TCP orientation and the target material impacted sensor performance. For all three sensors, the ADR percentage decreased as the angle of incidence increased. All three sensors had noticeably lower ADR scores for black polyester. The AKDK and OS0 express lower ADR results for denim, flannel, and black polyester at the largest angle whereas the D435i clearly

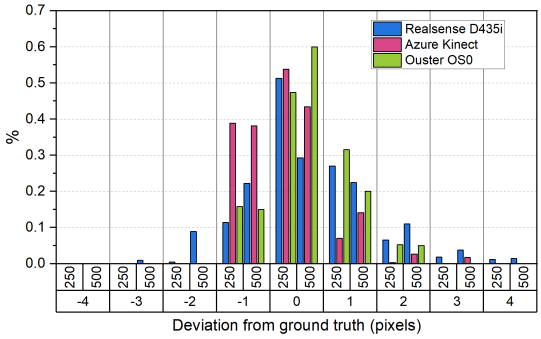


Fig. 9: Pixel spread for each sensor in dynamic edge precision testing

struggled most with black polyester throughout all tested angles. It is important to note that the ADR threshold chosen significantly impacted the scores across distances. Figure 7 shows the ADR performance using a threshold base on the average σ over all materials for each sensor.

E. Dynamic Path Tests

The first motion in the dynamic path tests was linear movement at a target. This path was the dynamic version of the static sphere depth and RRA experiments. Figure 10 shows that the D435i, AKDK and OS0 exhibited slightly raised μ errors with similar trends to static depth shots. The RRA

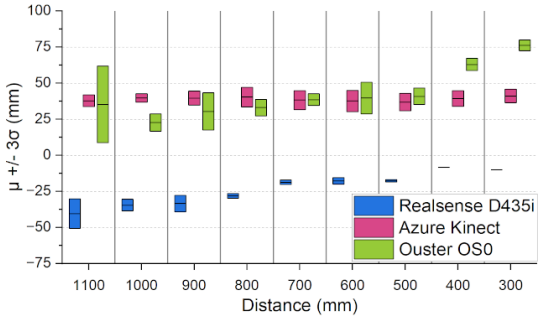


Fig. 10: Depth data for linear move towards the target

results in Fig. 11 show slightly degraded performance for the AKDK and the LiDAR. In contrast, the D435i presented a

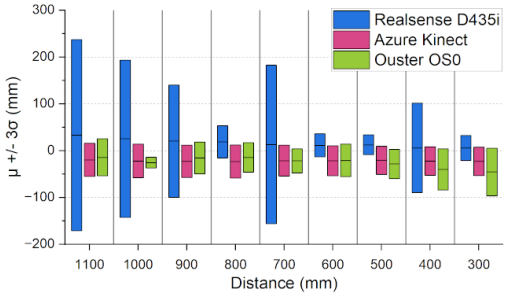


Fig. 11: RRA data for linear move towards the target

significant increase in RRA σ error and μ did not converge to 0 error as seen in the static experiment.

In path two, the target passes by the FOV linearly. In the case of LiDAR, the target is seen through the entire path, whereas the AKDK and D435i start and end with the foam sphere out of view. The Z distance is fixed in these tests, and the sensor travels along the x-axis. In Fig. 12, the captured frames were evenly distributed between when the sphere entered and exited the FOV of the sensors. In this test, the μ error for each sensor was visibly parabolic with the LiDAR exhibiting the highest error of 95mm in its second sample. The D435i, however, only generated a maximum depth μ error magnitude of 52mm. For RRA seen in Fig. 13, the D435i also demonstrated the least radius error compared to ground truth, and the LiDAR again exhibited the most error.

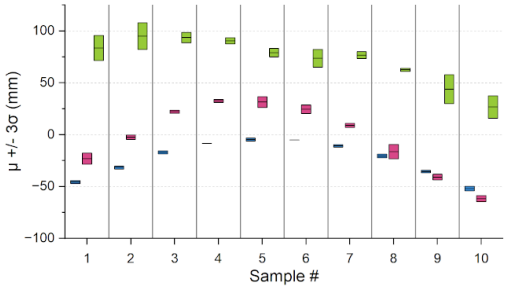


Fig. 12: Depth data for linear move across the target

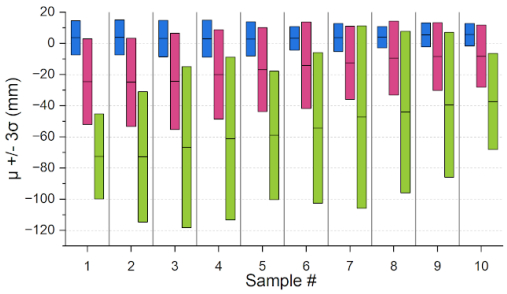


Fig. 13: RRA data for linear move across the target

The third motion was another sensor sweep across the target. In this motion, however, the robot took a radial path where the closest point to the target was set to be 300mm from the sphere surface. The particular trend of note in Fig. 14 is that the sensor behaviors in this radial motion created very similar error patterns to that of the linear motion. As seen in Fig. 15, the D435i exhibited the least RRA σ and μ error while the OS0 produced the most. The final motion consisted of a radial

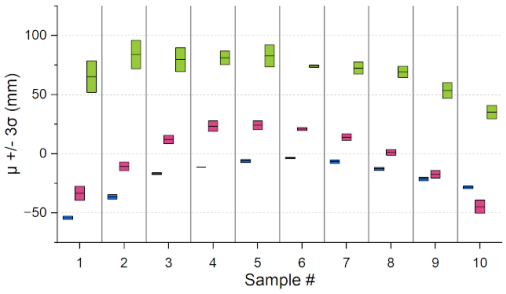


Fig. 14: Depth data for radial move across the target

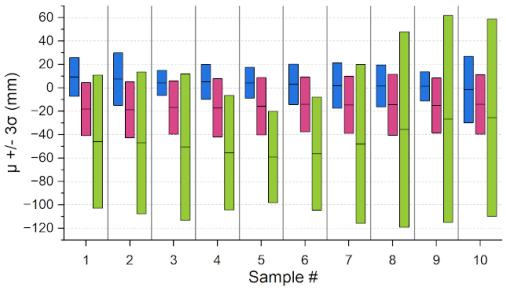


Fig. 15: RRA data for radial move across the target

move at the target. This motion generated the most depth and RRA error of the four paths for the ToF and stereo cameras. The samples in Fig. 16 were evenly distributed from full target entry until the target was entirely in the center of the frame. The depth error for the AKDK and LiDAR started close to zero and then increased to 40mm and 72mm, respectively. The D435i started near -100mm error and rose to less than 1mm of error. All three sensors ended their μ values very close to the final frames of motion path one, which lines up from a repeatability standpoint because both path 1 and 4 ended their paths with the foam sphere centered in the frame at 300mm away from the sensor mounted on the robot.

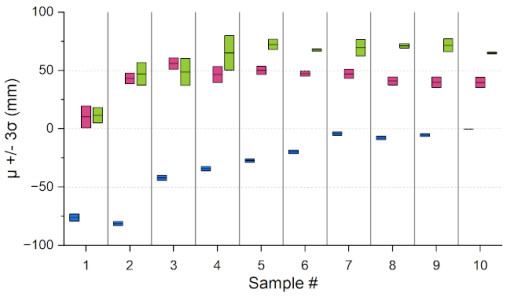


Fig. 16: Depth data for radial move towards the target

VI. Conclusion

This evaluation examined the performance of the D435i Realsense Camera, the AKDK ToF Camera, and the OS0-32

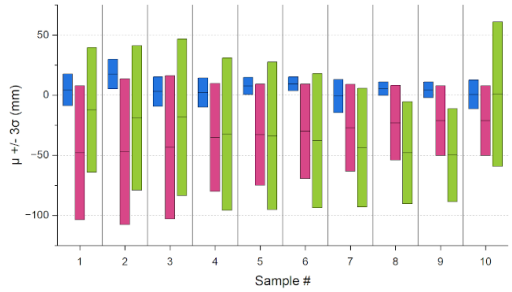


Fig. 17: RRA data for radial move towards the target

Lidar when exposed to robot paths and target types commonly seen in speed and separation monitoring applications. The data collected in these tests showed that the D435i has decent accuracy for the price when compared with the ToF and LiDAR sensors. Additionally, it was observed that the ToF camera exhibited the best flat depth μ and σ of the three sensors while the D435i generated the best RRA μ and σ. Though the AKDK is physically the largest module, the actual depth camera inside the AKDK is comparable in size to the D435i Realsense module. For cost, performance, and size, the ToF and Stereo cameras are promising sensing nodes for a point-rich SSM on industrial robots.

VII. Future Work

The D435i and AKDK performed well enough to pursue multi-sensor on-robot sensing modules. However, from an on-robot feasibility standpoint, using multiple Kinects on the TCP would significantly reduce the load capacity of the robot. Therefore, future CW ToF testing should be done using a smaller and lighter-weight module. The EVAL-ADTF3175 is a one-megapixel CW ToF depth camera module made by Analog Devices [2]. This sensor is smaller, has a comparable FOV to the D435i, and only requires a USB 3.1 compatible cable. Future tests should examine how well multiple sensors could work together. Multi-sensor integration could be crucial for combating occlusion caused by the robot seeing itself and general structures in the collaborative work environment. The performance of these point-rich on-robot perception systems will be compared with the existing on-robot 1-D ToF sensor rings described in [13].

Acknowledgment

The authors thank the Multi-Agent Bio-Robotics Lab (MABL) and the CM Collaborative Robotics Research (CMCR) Lab for their valuable inputs. The Authors would also like to thank Define Design Deploy Corp. for providing the Microsoft AKDK for the experiments performed in this paper.

References

[1] Odysseus Alexander Adamides, Anmol Saiprasad Modur, Shitij Kumar, and Ferat Sahin. A time-of-flight on-robot proximity sensing system to achieve human detection for collaborative robots. In *IEEE International Conference on Automation Science and Engineering*, volume 2019-Augus, pages 1230–1236. IEEE Computer Society, 8 2019.

[2] Analog Devices. ADTF3175 Datasheet and Product Info — Analog Devices, 2023.

[3] Charles Patrick C. Andres, Jason Patrick L. Hernandez, Lourdes T. Baldelomar, Christian Dior F. Martin, John Paul S. Cantor, Joycelyn P. Poblete, Jasmin D. Raca, and Ryan Rhay P. Vicerra. Tri-modal speed and separation monitoring technique using static-dynamic danger field implementation. *2018 IEEE 10th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management, HNICEM 2018*, 3 2019.

[4] Liliana Antão, João Reis, and Gil Gonçalves. Voxel-based Space Monitoring in Human-Robot Collaboration Environments. *IEEE International Conference on Emerging Technologies and Factory Automation, ETFA*, 2019-September:552–559, 9 2019.

[5] A. D. Brink. Thresholding of digital images using two-dimensional entropies. *Pattern Recognition*, 25(8):803–808, 8 1992.

[6] Christoph Byner, Björn Matthias, and Hao Ding. Dynamic speed and separation monitoring for collaborative robot applications – Concepts and performance. *Robotics and Computer-Integrated Manufacturing*, 58:239–252, 8 2019.

[7] Konrad P. Cop, Arne Peters, Bare L. Zagar, Daniel Hettegger, and Alois C. Knoll. New Metrics for Industrial Depth Sensors Evaluation for Precise Robotic Applications. *IEEE International Conference on Intelligent Robots and Systems*, pages 5350–5356, 2021.

[8] Urban B. Himmelsbach and Thomas M. Wendt. Built-In 360 Degree Separation Monitoring for Grippers on Robotic Manipulators in Human-Robot Collaboration. *Proceedings - 4th IEEE International Conference on Robotic Computing, IRC 2020*, pages 156–160, 11 2020.

[9] Urban B. Himmelsbach, Thomas M. Wendt, and Matthias Lai. Towards safe speed and separation monitoring in human-robot collaboration with 3D-time-of-flight cameras. *Proceedings - 2nd IEEE International Conference on Robotic Computing, IRC 2018*, 2018-Janua:197–200, 4 2018.

[10] Intel. Intel® RealSenseTM Product Family D400 Series, 2023.

[11] International Organization for Standardization. ISO/TS 15066:2016(en) Robots and robotic devices — Collaborative robots, 2016.

[12] Kuka. LBR iiwa — KUKA AG, 2023.

[13] Shitij Kumar, Celal Savur, and Ferat Sahin. Dynamic Awareness of an Industrial Robotic Arm Using Time-of-Flight Laser-Ranging Sensors. *Proceedings - 2018 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2018*, pages 2850–2857, 1 2019.

[14] Shitij Kumar, Celal Savur, and Ferat Sahin. Survey of Human-Robot Collaboration in Industrial Settings: Awareness, Intelligence, and Compliance. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 51(1):280–297, 1 2021.

[15] Bakir Lacevic and Paolo Rocco. Kinetostatic danger field - A novel safety assessment for human-robot interaction. *IEEE/RSJ 2010 International Conference on Intelligent Robots and Systems, IROS 2010 - Conference Proceedings*, pages 2169–2174, 2010.

[16] Marco Esposito, O'Flaherty Rowland, Yihui Li, Salvo Virga, Ravi Joshi, Robert Haschke, Hui Xiao, and Berkay Cakal. GitHub - IFL-CAMP/easy_handeye: Automated, hardware-independent Hand-Eye Calibration, 2022.

[17] Microsoft. Azure Kinect DK hardware specifications — Microsoft Learn, 2022.

[18] Optitrack. OptiTrack - Flex 13 - An affordable motion capture camera, 2023.

[19] OptiTrack. OptiTrack - Motive - In Depth, 2023.

[20] Ouster Inc. OS0 Ultra-Wide View High-Resolution Imaging Lidar, 2023.

[21] Adriano Scibilia, Marcello Valori, Nicola Pedrocchi, Irene Fassi, Sebastian Herbster, Roland Behrens, Jose Saenz, Alice Magisson, Catherine Bidard, Morten Kuhnrich, Aske Bach Lassen, and Kurt Nielsen. Analysis of Interlaboratory Safety Related Tests in Power and Force Limited Collaborative Robots. *IEEE Access*, 9:80873–80882, 2021.

[22] Petr Svarny, Michael Tesar, Jan Kristof Behrens, and Matej Hoffmann. Safe physical HRI: Toward a unified treatment of speed and separation monitoring together with power and force limiting. *IEEE International Conference on Intelligent Robots and Systems*, pages 7580–7587, 11 2019.

[23] Barnaba Ubezio, Christian Schoffmann, Lucas Wohlhart, Stephan Mulbacher-Karrer, Hubert Zangl, and Michael Hofbaur. Radar Based Target Tracking and Classification for Efficient Robot Speed Control in Fenceless Environments. *IEEE International Conference on Intelligent Robots and Systems*, pages 799–806, 2021.

[24] Universal Robotics. Universal Robots e-Series User Manual, 2022.