Towards the development of standards and performance metrics for 3D imaging systems

Prem Rachakonda[a], Kamel Saidi[a], Marek Franaszek[a], Wesley Rhodes[a], Helen Qiao[a],
John Sweetser[b], Armin Khatoonabadi[c], David Dechow[d]
[a]National Institute of Standards and Technology, Gaithersburg, MD, USA; [b]RealSense Group, Intel Corporation, Santa Clara, CA, USA; [c]Apera AI, Vancouver, BC, Canada;
[d]Machine Vision Source LLC, Salisbury, NC, USA;

ABSTRACT

NIST and multiple industrial stakeholders are leading and supporting multiple efforts to develop standards for 3D imaging systems for manufacturing automation applications. Many manufacturers specify the performance of their sensors in non-standard ways and offer no method to verify those parameters independently. The standards are being developed under the auspices of the ASTM E57 committee on 3D Imaging Systems. They are meant to produce a) standards for measuring the performance of 3D imaging systems, b) standards for bin-picking vision systems, and c) guidelines for the selection of 3D imaging systems. This work presents the status of four work items.

Keywords: standards, sensors, 3D imaging, manufacturing automation, performance metrics

1 INTRODUCTION

The Intelligent Systems Division (ISD) at the National Institute of Standards and Technology (NIST) develops, advances, and deploys measurement science and standards to speed developing, adopting, and integrating leading-edge intelligent technologies to advance U.S. manufacturing performance. One of the programmatic focus areas of NIST/ISD is sensors and perception systems for robotic applications. As part of this focus area, NIST is performing research and helping the industry develop documentary standards. It also works with the ASTM subcommittee E57.23 on Industrial 3D Machine Vision Systems.

NIST staff have been the founding members of the ASTM committee E57 on 3D imaging systems (in 2006), which has published several standards. The E57.23 subcommittee addresses the performance evaluation of 3D imaging systems for manufacturing automation and vision-guided robotics. The applications of 3D imaging systems extend to other areas, such as autonomous vehicles, 3D reconstruction, and entertainment. NIST conducted a workshop in 2019[1] to seek input from industry leaders, practitioners, and researchers worldwide on standards for 3D perception systems for robotic assembly applications.

After the 2019 workshop, NIST selected four topic areas with the most significant impact, and four ASTM work items were formed to address them. The four work items are:

1. WK72962 - Standard Test Method for Measuring the Performance of a 3D Perception System Across the Specified Field-of-View (FOV).
2. WK73176 - Standard Test Methods for Determination of a 3D Perception Systems Point Wise Spatial Resolution.
3. WK78941 - Standard Test Method for Performance of Machine Vision Systems for Bin-picking Systems.
4. WK81247 - Standard Guide for Selection of 3D Vision Technologies for Industrial Applications.

This paper will describe the activities of these standards' working groups.

2 3D IMAGING SYSTEMS

A 3D imaging system is a non-contact measurement instrument used to produce a 3D representation (for example, a point cloud) of an object or a site [2]. They may measure the dimensions, position, orientation, and velocity of an object or a surface. The 3D imaging systems being considered for this work are primarily optical instruments that use either the principles of stereoscopy, time-of-flight, or triangulation. These include sensors such as lidars, structured light systems, and active/passive stereo systems that use pseudo-random patterns.

Of the numerous varieties of sensors, this project procured sensors that are typically used for manufacturing automation applications, listed in Table 1.

Table 1: Relevant specifications of the sensors used in this work [3]

Sensor names[*,†]	Technology	Shutter Type	Frames per sec.	Light Source Wavelength	Depth Image # of Pixels	Depth Working Range (m)
Z0	Structured light	N/A	N/A	White light	1920 × 1200	0.30 – 1.3
Z1	Structured light	N/A	N/A	White light	1920 × 1080	0.70 – 1.5
Z2	Structured light	N/A	N/A	White light	1944 × 1200	0.30 – 1.3
E2	Active stereo	Global	10	465 nm	1280 × 1024	0.50 – 3.0
D3	Active stereo	Global	30-90	≈ 850 nm	1280 × 720	0.60 – 6.0
L4	Time-of-flight	Global	30	≈ 850 nm	1024 × 768	0.25 – 9.0
K5	Time-of-flight	Rolling	30	≈ 860 nm	1024 × 1024	0.25 – 5.5

3 WK72962: STANDARD TEST METHOD FOR MEASURING THE PERFORMANCE OF A 3D PERCEPTION SYSTEM ACROSS THE SPECIFIED FIELD-OF-VIEW (FOV).

The ASTM work item WK72962 was constituted to work on defining metrics for a 3D imaging system's performance across the FOV and scoped to include the following five distinct metrics:

a) Depth error: The difference between the depth value reported by the 3D Imaging system and the corresponding reference value.
b) Fill ratio: The ratio of valid pixels in a selected region of a depth image to the total number of pixels in the selected region, stated as a percentage.
c) Pixel temporal noise: The standard deviation of the depth values corresponding to each pixel over a certain number of depth frames captured by the system under test (SUT).
d) Frame temporal noise: The standard deviation of the mean depth error of each depth frame captured by the SUT.
e) Root-mean-square-error (RMSE): The root-mean-square of the orthogonal distances of the points on the target from a best-fit plane to the points.

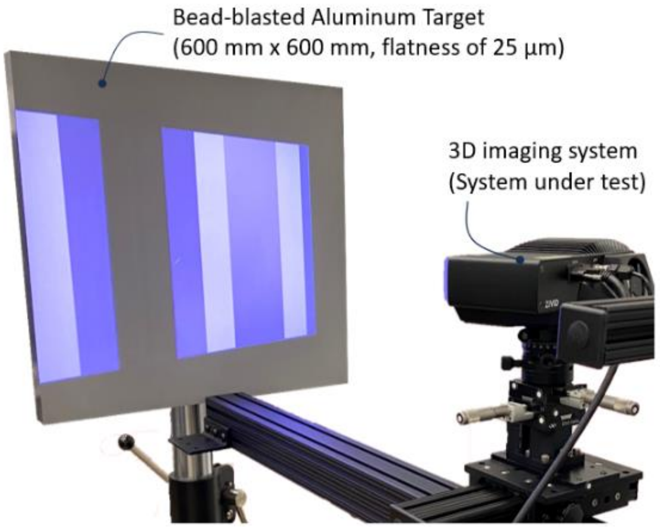


Figure 1: A setup with a target and the system under test mounted on an apparatus to allow alignment and positioning.

Depth is defined as the distance from a reference plane of the 3D imaging system to a point on the surface of an object within the system's FOV. Of the five metrics being developed, the tests for depth error are the most challenging. One of the key reasons is the need to physically identify the reference plane where the depth value is zero. Such a reference plane is typically inside the sensor and inaccessible to the end-user. Any depth error estimate must consider the uncertainty in measuring the reference plane. For some of the sensors listed in Table 1, the manufacturer provides the distance of the sensor's reference plane from the front or the back of the sensor's enclosure.

---

[*] Sensor names are intentionally anonymized.

[†] 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.

3.1 Target flatness

One of the critical considerations for developing a standard is to reduce the barriers to implementing the standard by the end-user. One such challenge is to find a target of low flatness and large dimensions to test the sensor in its entire FOV. Per ASME Y14.5-2018, flatness specifies a tolerance zone defined by two parallel planes within which the surface, or derived median plane, shall lie.

The system under test (SUT) may provide multiple points on the target surface (typically 100's to 1000's; see Table 2). The SUT provides point cloud data spread evenly over the target's surface. Depending on the method of measuring ground truth using a reference instrument (RI), the reference depth measurements may not have the same even spread or point density over the region of interest. This is because an RI may only measure at the corners of the target and sometimes only at the target's center, and these may not be on the median plane of the target. A tighter flatness specification is necessary to have a lower uncertainty in the measurement of reference depth. Figure 2 shows point clouds of three identical square targets (≈ 200 mm × 200 mm × 13 mm)as measured on a coordinate measuring machine (CMM) and their flatness indicated.

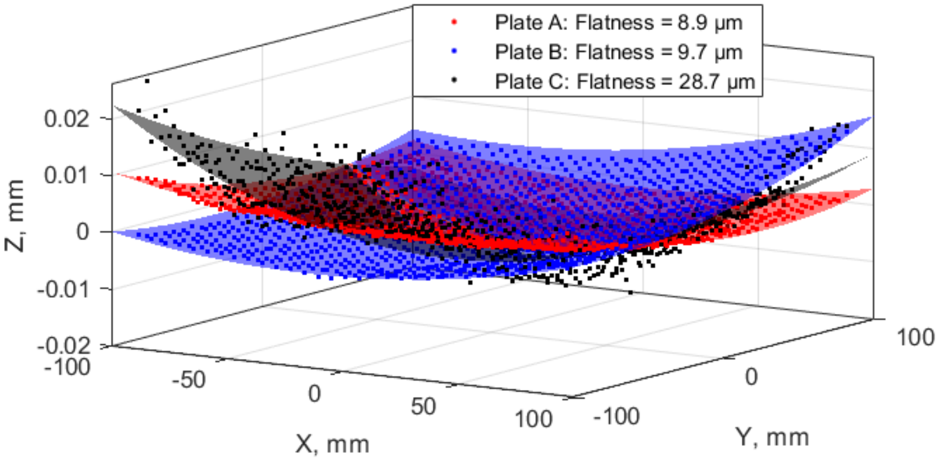


Figure 2: Point cloud data on three flat targets measured on a CMM and their flatness values (axes not to scale)

3.2 Target parallelism to the sensor reference plane

One of the other key considerations for this work is the alignment of the target parallel to the sensor's reference plane. Several methods are being considered for this process. However, it should be noted that perfect alignment of a target may not be possible, and in this work, an angular deviation of ±3° with respect to the SUT was achieved with relative ease. Any greater angular misalignment will result in a more significant depth error and must be accounted for in the uncertainty budget.

4 WK73176: STANDARD TEST METHODS FOR DETERMINATION OF A 3D PERCEPTION SYSTEMS POINT WISE SPATIAL RESOLUTION

Depth resolution is the smallest change in depth that causes a perceptible change in the corresponding measured or derived depth. Note that perceptible change is detected by an algorithm, not necessarily by the human eye. Rachakonda et al. [4] describe the procedure and apparatus to calculate depth resolution. This method's underlying premise is that the ability of a system to resolve a feature is dependent on a) depth noise, b) quantization of data that are observed in the depth direction, and c) point density. Essentially, larger depth noise or data quantization results in lower resolution. Table 2 shows various parameters calculated for each of the sensors from Table 1, including the resolution described by [4].

Unlike the test methods described in Section 3, depth resolution is less sensitive to target flatness. This is because depth resolution refers to the change in the depth along the depth direction (along Z-axis in this work). Any flatness variations will have minimal influence on the quantization error and on the variation in the Z-coordinate of the centroid at adjacent spatial locations in close proximity along the Z-axis and, thereby, on the resolution. However, surfaces with significant spatial variations (bumpy surfaces) and large angles of the target relative to the sensor can increase the errors in depth measurement. For these reasons, one of the recommendations of this work is to limit the target flatness to 1 mm and the target's parallelism to ±3°.

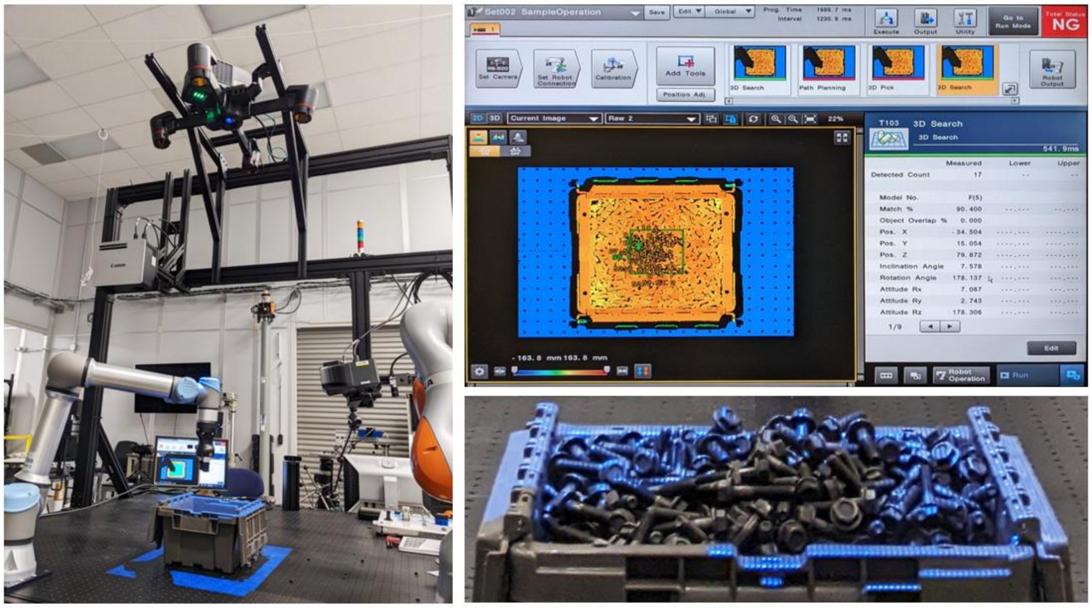
A new method to calculate the uncertainty of these measurements was also developed. This method propagates uncertainty by performing the Monte Carlo simulation of the quantization error, assuming a uniform distribution, and performing a bootstrap sampling of the Z-coordinates of the target region's centroids. The resulting expanded uncertainties of these values are given in Table 2. The data was obtained on a setup similar to the setup shown in Figure 1 on a target with flatness <10 µm and using 20 frames of data from each sensor. A detailed description of the values in Table 2 is provided in [4].

Table 2: Resolution and related parameters of various sensors on a target area of 50 mm × 50 mm

Sensor	Dist. (mm)	Avg. # of pts (P)	Quantization error (mm)	1σ of plane Z-coordinate of centroids (mm)	Resolution (mm)	Expanded Uncertainty (mm)
Z0	542.9	65736	0.0002	0.001	0.004	0.001
Z1	807.3	29043	0.0002	0.016	0.049	0.016
Z2	692.5	16378	0.0002	0.009	0.026	0.005
E2	633.3	13222	0.2112	0.023	0.194	0.033
D3	507.9	501	0.2500	0.441	1.323	0.364
L4	612.9	342	0.2500	0.566	1.688	0.261
K5	398.3	3772	1.0000	0.053	0.869	0.165

5 WK78941: STANDARD TEST METHOD FOR PERFORMANCE OF MACHINE VISION SYSTEMS FOR BIN-PICKING SYSTEMS

This work item defines a quantitative test procedure for measuring different aspects of a 3D imaging system's performance for robotic picking of objects with known geometries. Presently, this work item focuses on defining and measuring key performance metrics such as vision cycle time, robot cycle time, and pose accuracy of the vision system. To this end, NIST/ISD established a testbed with five different commercial bin-picking systems and one custom-developed system (see Figure 3).



T103 3D Search

3D Search

541.9ms

	Measured	Lower	Upper
Detected Count	17	--	--
Model No.	F(5)		
Match %	90.400	--.---	--.---
Object Overlap %	0.000		
Pos. X	-34.504	----.---	----.---
Pos. Y	15.054	----.---	----.---
Pos. Z	79.872	----.---	----.---
Inclination Angle	7.578	---.---	---.---
Rotation Angle	178.137	---.---	---.---
Attitude Rx	7.067	---.---	---.---
Attitude Ry	2.743	---.---	---.---
Attitude Rz	178.306	---.---	---.---

1/9

Figure 3: The NIST bin-picking testbed

One of the critical challenges in this work is to separate and measure the cycle times described earlier. One of the reasons is that most commercial systems do not offer information on the start and end of these cycle times. Another reason is that these cycle times depend on the object texture/color/finish, gripper design, and stability of the objects. Further, it is unclear how to quantify the surface finish (such as gloss or roughness) relevant to 3D imaging systems. This is because some rough surfaces can have varying gloss (shiny/dull).

6 WK81247: GUIDE TO THE SELECTION OF 3D IMAGING SYSTEMS

The ASTM work item WK81247 is focused on developing a guide to provide users with information on best practices in the selection of typical commercial 3D vision technologies for robotic bin-picking. This guide is intended to help the end-user select, integrate, and execute 3D vision systems technologies to improve suitable manufacturing operations and specific requirements. This guide will aid the users of 3D imaging technologies in discussions with potential suppliers and systems integrators.

The guide is presently focusing on four primary sections:

a) Landscape: This section provides a general overview of the commercially available methods for creating, representing, and processing 3D images, including hardware, software, and data formats typically used in these commercial systems.
b) Application Needs: This section guides users in understanding the specific needs of their application, including their unique requirements, environmental conditions, certain commercial considerations, and other essential factors.
c) 3D Imaging System Metrics: This section provides a basic understanding of the key metrics of various 3D imaging technologies and how these metrics affect the performance of a bin-picking solution.
d) Converting from Metrics to a Solution: This section guides users on converting their application needs into metrics and selecting certain 3D imaging systems based on these metrics.

7 SUMMARY & FUTURE DIRECTIONS

NIST and multiple stakeholders are working on developing multiple standards for 3D imaging systems in collaboration with the industry. The purpose of these standards is to reduce confusion among end-users, integrators, and researchers about the performance of these systems, enable one-to-one comparisons, and help end-users make informed choices.

The four work items described in this paper are led by industry experts, with support from NIST. NIST's Intelligent System Division is equipped with state-of-the-art instrumentation for ground truth measurements, various 3D sensors, and an environmentally stable laboratory. NIST also has access to other calibration laboratories, fabrication facilities, and experts in various fields of science and engineering. These four work items are in multiple stages of progress, with WK72962 and WK73176 (on depth-related metrics) undergoing a second review of the draft documents and WK78941 and WK81247 in the draft development stage. Tests are planned to realize many methods described in WK72962 and WK78941. More industry input will accelerate its development and adoption. Such standards are expected to improve confidence in these systems, improve industrial competitiveness, and enhance US commerce.

ACKNOWLEDGMENTS

The authors would like to thank members of the ASTM E57 work items for their valuable contributions and feedback that went into this work.

REFERENCES

---

[1] Saidi, K. , Cheok, G. , Qiao, G. , Horst, J. and Franaszek, M. (2020), Proceedings of the ASTM E57 Workshop on Standards for 3D Perception Systems for Robotic Assembly Applications December 2 &3, 2019, Advanced Manufacturing Series (NIST AMS), National Institute of Standards and Technology, Gaithersburg, MD, [online], https://doi.org/10.6028/NIST.AMS.100-33

[2] Standard Terminology for Three-Dimensional (3D) Imaging Systems https://www.astm.org/e2544-11ar19e01.html

[3] A Comprehensive List of 3D Sensors Commonly Leveraged in ROS Development https://rosindustrial.org/3d-camera-survey

[4] Rachakonda, P., Cheok, G., Franaszek, M., Saidi, K., "Evaluating the depth resolution of 3D sensors for manufacturing automation applications", Measurement, Volume 220, 2023, https://doi.org/10.1016/j.measurement.2023.113307