NIST Advanced Manufacturing Series 100-33

Proceedings of the ASTM E57 Workshop on Standards for 3D Perception Systems for Robotic Assembly Applications December 2 – 3, 2019

Kamel Saidi
Geraldine Cheok
Helen Qiao
John Horst
Marek Franaszek

This publication is available free of charge from:
https://doi.org/10.6028/NIST.AMS.100-33

National Institute of Standards and Technology
U.S. Department of Commerce

NIST Advanced Manufacturing Series 100-33

Proceedings of the ASTM E57 Workshop on Standards for 3D Perception Systems for Robotic Assembly Applications December 2 – 3, 2019

Kamel Saidi
Geraldine Cheok
Helen Qiao
John Horst
Marek Franaszek
Engineering Laboratory
National Institute of Standards and Technology
Gaithersburg, Maryland 20899

This publication is available free of charge from:
https://doi.org/10.6028/NIST.AMS.100-33

April 2020

U.S. Department of Commerce
Wilbur L. Ross, Jr., Secretary

National Institute of Standards and Technology
Walter Copan, NIST Director and Undersecretary of Commerce for Standards and Technology

National Institute of Standards and Technology Advanced Manufacturing Series 100-33 Natl. Inst. Stand. Technol. Adv. Man. Ser. 100-33, 103 pages (April 2020)

This publication is available free of charge from: https://doi.org/10.6028/NIST.AMS.100-33

Abstract

A two-day workshop was held at the National Institute of Standards and Technology (NIST) in Gaithersburg, MD on Dec. 2 – 3, 2019 on the subject of defining and prioritizing standards needed for three-dimensional (3D) robotic perception systems for part localization, part recognition, part capture, and complex assembly. The event had 28 attendees, including 10 NIST personnel. The work that this workshop was a part of seeks to benefit both vendors and users in the 3D perception industry through the development of new standards for a variety of use cases. Key results from the workshop were: 1) a ranked list of 39 standards needed, 2) six of the highest ranked standards were developed into work items within ASTM International, and 3) several of the attendees from industry agreed to champion some of the six work items. This document is a report on the various presentations, activities, decisions, and interactions that happened during and prior to the workshop.

Key words

3D machine vision, vision standards, robotic perception, robots, robotic assembly, 3D imaging, standard artifacts, manufacturing

Table of Contents

List of Tables

1 Introduction/Background

On Dec. 2-3, 2019, the National Institute of Standards and Technology[1] (NIST) and ASTM International Committee E57 on three-dimensional (3D) Imaging Systems co-sponsored a workshop on Standards for 3D Robotic Perception Systems and Robotic Assembly Applications.

1.1 NIST Program/Project background

The workshop was organized by members of the Perception Performance of Robotic Systems (PPRS) project[2] and was held at NIST's headquarters in Gaithersburg, Maryland. The PPRS project is part of the Measurement Science for Manufacturing Robotics program[3] under NIST's Engineering Laboratory's[4] Intelligent Systems Division[5].

The PPRS project's objective is to develop measurement science for sensing and perception system performance characterization to reduce the risk related to the adoption of new technologies and to advance the agility, safety, and productivity of collaborative industrial and mobile robots. Performance standards can help achieve this objective, and the project is actively attempting to understand the standards needs of the industry.

1.2 Pre-workshop meetings and results

In April 2019, the PPRS project and ASTM E57 co-sponsored a meeting at the 2019 Automate Show in Chicago, Illinois that was attended by representatives from the 3D perception industry. The purpose of this meeting was to kickoff the effort of developing a roadmap of standards that are needed for 3D perception systems that can be used for robotic assembly. Fifteen participants (including 10 from industry and academia) conducted a brainstorming exercise in which 49 ideas for needed standards were identified and grouped into related categories (see Table 1). The initial grouping in Table 1 was based on perceived similarity and arbitrarily labeled from A to T.

Following the kickoff meeting, an ad hoc working group (AC475) was created under ASTM E57 in order to refine and consolidate the list of 49 ideas of needed standards. The working group included representatives from several major sensor and robot manufacturers as well as users of sensors from industry, academia, and government. The working group held 7 virtual meetings between June and November of 2019 during which the original list of 49 ideas were further developed with regards to intent and purpose. The ideas were then grouped and condensed into 27 original ideas and 3 new ideas, proposed during the virtual meetings, that were discussed during the workshop. These ideas are listed in Table 2 and were discussed during the workshop.

[1] https://www.nist.gov/
[2] https://www.nist.gov/programs-projects/perception-performance-robotic-systems
[3] https://www.nist.gov/programs-projects/measurement-science-manufacturing-robotics
[4] https://www.nist.gov/el
[5] https://www.nist.gov/el/intelligent-systems-division-73500

Table 1. Original 49 Proposed Standards from the April 2019 Meeting.

Proposed Ideas The following ideas were proposed during the April 10th meeting at Automate 2019 and are presented verbatim as originally written.	**Initial Grouping** The letters in this column indicate the ideas that were grouped together during the meeting.
Classification of "interpretation" methods, semantic, process-based, etc.	A
Human Recognition	B
Predictive Movement	B
Angular Resolution (changes when further away)	C
Capturing xy versus radial (resolution)	C
Data (xyz) resolution	C
Define resolutions	C
Measuring depth resolution	C
X, Y resolution + Z (error)	C
Minimum set of classification/interpretation types for a system to be considered "able to perceive"	D
Reference objects or artifacts to use for benchmarking (a la YCB set – http://www.ycbbenchmarks.com/)	D
Standard test(s) & artifacts	D
Test/calibration targets	D
Test method as part of a system or just the sensor? Does the 'sensor' include the comms and display?	E
Standards for output data format (dense versus sparse colorized, intensity, etc.)	F
Ability to resolve sharp edges and corners	G
Error against traceable targets	G
Standardized target recognition algorithms	G
Changes in metrics throughout the field of view (e.g., depth error varies with distance)	H
Single part versus cluttered scene 6DOF accuracy	I
Inter-operability	J
Error Sensing	K
Outlier / Error Rate	K
Measurement volume (FOV, MR, CD, …)	L
Speeds and the effects on uncertainty	M
XYZ linearity	N
Cycle time	O

Proposed Ideas The following ideas were proposed during the April 10th meeting at Automate 2019 and are presented verbatim as originally written.	**Initial Grouping** The letters in this column indicate the ideas that were grouped together during the meeting.
Auto- ISO 24262 on chip? Functional safety with or without software	P
Power over FOV = eye safe	P
Change in depth resolution between 10 klx – 100 klx ambient	Q
Interference in sensor data due to implementation on a system	Q
Interfering wavelengths. Interference with itself?	Q
Quantify error due to occlusions	Q
Altering ambient [light]	R
Ambient conditions	R
Temperature range and effects	R
Temperature stability	R
Vibration Specifications	R
Error from specular reflection	S
Part reflectivity test	S
Part shape effects on accuracy. E.g., interreflections, concave parts	S
Part surface detection accuracy depth + color built in AI	S
Reflectivity of object versus depth error	S
Robustness to surface reflectivity	S
Global shutter at time of capture versus receipt of image	T
Latency	T
Latency versus integration versus frame rate	T
Real-time versus latency	T
Time sync IEEE 1988	T

Table 2. List of 30 Proposed Standards for Discussion at the Workshop.

Idea	Category	Description
Pointcloud XYZ resolution	Resolution	Standards for evaluating the smallest measurements that a system can achieve in the X, Y, and Z directions for 3D perception systems that produce pointclouds from a single sensor or multiple sensors.
Depth map XYZ resolution	Resolution	Standards for evaluating the smallest measurements that a system can achieve in the X, Y, and Z directions for 3D perception systems that produce depth maps.
2D image XY resolution	Resolution	Standards for evaluating the smallest measurements that a system can achieve in the X and Y directions for 3D perception systems that produce two-dimensional (2D) images.
Part position resolution	Resolution	Standards for evaluating the smallest changes of a part's position along the X, Y, and Z axes that a 3D perception system can measure.
Part orientation resolution	Resolution	Standards for evaluating the smallest changes of a part's orientation about the X, Y, and Z axes that a 3D perception system can measure.
System-to-part suitability	System Suitability	Standards to determine whether a 3D perception system is appropriate for determining the pose of a part for a particular application, e.g., is a particular system useful for small, metal automotive parts?)
Standard reference objects or artifacts	Standard Reference Objects	Standards describing reference objects that can be used for benchmarking and/or calibrating a 3D perception system's performance (e.g., interreflections, concave vs. convex parts, curved vs. planar surfaces, etc.).
Ability to resolve geometric features	Standard Reference Objects	Standards for measuring a 3D perception system's ability to resolve geometric features (e.g., edges and corners) on standard reference objects.
Error against traceable targets	Standard Reference Objects	Standards for using standard reference objects to evaluate a 3D perception system's errors.
Changes in performance throughout a perception system's field-of-view (FOV)	FOV	Standards for measuring a 3D perception system's performance throughout its FOV.
Measurement volume specification/verification	FOV	Standards for measuring a 3D perception system's measurement volume (FOV, measurement range, calibrated distance, standoff distance, etc.)

Idea	Category	Description
Interoperability	Interoperability	Standard protocols, data formats, or interfaces to allow sensors from different vendors to work with software/robots from different vendors.
Output quality	Self-Diagnostics	Standards for measuring a 3D perception system's ability to quantify the quality of the output (e.g., values for different types of errors, confidence in 6DOF pose, false positives, measurement dispersion over time, etc.).
Dynamic performance	Dynamic Performance	Standards for measuring the effects of sensor (or object) motion on a 3D perception system's part-pose measurement performance (e.g., ASTM E3064).
XYZ linearity	Linearity	Standards for measuring how linear a 3D perception system's measurements are in X, Y, and Z.
Functional safety	Safety	Standards for evaluating a 3D perception system's functional safety (i.e., its ability to properly handle likely human errors, hardware failures and operational/environmental stress - e.g., ISO 26262).
Eye safety over FOV	Safety	Standards for measuring the eye safety of a sensor's active illumination across its entire FOV.
Ambient conditions	Application Conditions	Standards for measuring the effects of changes in ambient conditions (lighting, temperature, humidity, vibrations, EMF interference, background specular reflections, etc.) on the 3D perception system's part-pose measurement performance.
Performance due to cluttered versus uncluttered scenes	Application Conditions	Standards for measuring a 3D perception system's ability to measure the 6DOF pose of a single part presented alone vs. a part presented within a cluttered environment.
Performance due to occlusions	Application Conditions	Standards for measuring the effects of part occlusion (self-occlusions or occlusions by other parts) on the 3D perception system's part-pose measurement performance.
Temperature stability	Drift	Standards for measuring the effects of changes in a 3D sensor's internal temperature on the 3D perception system's part-pose measurement performance.
Performance due to part material properties	Surface and Material Properties	Standards for measuring the effects of different part material properties on the 3D perception system's part-pose measurement performance (e.g., effects of light penetration).

Idea	**Category**	**Description**
Performance due to part surface properties	Surface and Material Properties	Standards for measuring the effects of different part surface properties on the 3D perception system's part-pose measurement performance (e.g., diffuse vs. specular reflections, reflectance, etc.).
Part reflectance	Surface and Material Properties	Standards for measuring part reflectance (e.g., parts with curved surfaces, multifaceted parts, parts with multiple reflectivities, etc.)
Latency	Latency and Timing	Standards for measuring the time between when a perception system is commanded to take a measurement and when a usable measurement is available to other systems, with possible definitions for "integration time," "frame rate," and "real-time" (e.g., ASTM 3124-17).
Time synchronization	Latency and Timing	Standards for measuring the time synchronization between different 3D sensors or systems (e.g., IEEE 1588).
Cycle time	Latency and Timing	Standards for measuring the time for a robotic system to estimate the 6DOF pose of a part, grip the part, and deliver the part to its final destination. (E.g., "cycle time" could be defined as the time it takes between the command to the 3D perception system to measure the 6DOF pose of a part until the pose is available for the robot to use - or until the robot acquires the part).
New ideas		
Static performance	Static Performance	Standards for evaluating a 3D perception system's static part-pose measurement performance (e.g., ASTM E2919).
Depth error	Error	Standards for evaluating a 3D perception system's depth error.
Bit precision resolution	Resolution	Standards for measuring a 3D perception system's ability to define the precision of the data

1.3 Purpose and objectives of the workshop

The purpose of the workshop was to bring together stakeholders in 3D perception systems (vendors/manufacturers, users, researchers, etc.) in order to:

1. Learn about the challenges, barriers, and solutions to implementing 3D perception systems for robotic applications;
2. Develop a roadmap of consensus standards needed for 3D perception systems; and
3. Identify high-priority standards for the manufacturing industry and organize ASTM task groups to develop those standards.

2 Workshop Description

The workshop was held over a period of two days on December 2nd and 3rd, 2019 at NIST in Gaithersburg, Maryland.

2.1 Program

The workshop program included seven technical presentations, three work sessions, several laboratory tours, and a panel discussion. The workshop program is presented in Table 3.

Table 3. Workshop Program

Dec. 2, 2019 – Day 1	
07:30 - 08:00	Arrival at NIST and Visitor Center Registration
08:00 - 08:15	Welcome
08:15 - 08:30	Introductions
08:30 - 08:50	NIST Overview and Workshop Motivation – Elena Messina & Kamel Saidi, NIST
08:50 - 09:20	Technical Presentation 1 – Remus Boca, ABB
09:20 - 09:50	Technical Presentation 2 – Miguel Saez, General Motors
09:50 - 10:05	Break
10:05 - 12:00	Work Session 1
12:00 - 13:00	Lunch
13:00 - 14:30	Lab Tours
14:30 - 15:00	Technical Presentation 3 – Michele Pratusevich, Root AI
15:00 - 15:30	Technical Presentation 4 – John Sweetser, Intel Corp.
15:30 - 15:45	Break
15:45 - 17:45	Work Session 2
17:45 - 18:00	Summary of Work Session 2
18:30 - 20:00	Group Dinner
Dec. 3, 2019 – Day 2	
08:00 - 08:15	Summary of Day 1
08:15 - 08:45	Technical Presentation 5 – Song Zhang, Purdue University
08:45 - 09:15	Technical Presentation 6 – Joseph Schornak, Southwest Research Institute
09:15 - 09:30	Break
09:30 - 11:15	Work Session 3
11:15 - 11:30	Break
11:30 - 12:00	Technical Presentation 7 – Jared Glover, Capsen Robotics
12:00 - 13:00	Lunch
13:00 - 15:00	Lab Tours
15:00 - 15:30	Summary of Work Session 3
15:30 - 15:45	Break
15:45 - 17:30	Panel Discussion
17:30	Adjourn

2.2 Participants

An effort was made to involve many 3D perception vendors, end users, and researchers in the workshop. Notifications of the workshop were sent to Lidar News, Quality Magazine, and the Collaborative Robots, Advanced Vision & AI (CRAV) conference. Advertisements were posted at the 2019 International Conference on

Intelligent Robots and Systems (IROS) conference. Personal communications were made to various researchers via email and to vendors at tradeshows and conferences.

Seventeen non-NIST participants (and ten from NIST) attended the workshop. The following organizations were represented:

- ABB
- Airy3D
- Brown University
- CapSen Robotics
- General Motors
- Intel RealSense
- May Solutions
- National Research Council of Canada
- Purdue University
- Root AI
- Sense Photonics
- Southwest Research Institute
- U.S. Postal Service
- Visio Nerf
- X-wave Innovations

A list of participants is given in Appendix A.

2.3 Technical Presentations

All of the technical presentation slides may be found in Appendix B. The technical presentation titles, author names and bios, and abstracts are provided below.

2.3.1 Technical Presentation 1

Title: Perception challenges for industrial applications

Author: Remus Boca (ABB)

Author Bio: Remus Boca joined ABB Corporate Research Center in 2010. He is a Senior Principal Scientist focusing on computer vision, sensing, perception, robotics and autonomy for industrial equipment and machines. He designs and implements strategies for machine perception and visual cognition targeting a wide range of ABB applications across different industrial segments such as robotics, shipyards, metallurgy, mining, electrical equipment, food & beverage, logistic and warehouse.

Prior to joining ABB, Remus worked at Braintech Inc as a Senior Robotic Vision Scientist on integrating perception solutions with industrial robots. He has a PhD, MS and bachelor's degrees in Industrial Robotics and Automation from University Politehnica Bucharest, Romania.

Abstract: As the world moves towards autonomy, the sensing and perception are becoming more important if not necessary. Industrial applications have their own challenges as they operate in possible harsh environments, they require continuous and robust operation, need to accommodate unstructured environments, determine a wide range of states and unexpected events. This talk presents perception needs and challenges across many industries such as ports, mining, industrial equipment inspection, logistic and robotics.

2.3.2 Technical Presentation 2

Title: Robotic Assembly: Challenges and Opportunities in the Automotive Industry

Author: Miguel Saez (General Motors)

Author Bio: Dr. Miguel Saez is currently a researcher for General Motors Research and Development, Manufacturing Systems Research Lab in Warren, Michigan. In his current role, he develops novel industrial robotics and automation solutions to advance the technology used for manufacturing electric vehicles. He holds a bachelor's degree in Mechanical Engineering from La Universidad del Zulia, Venezuela and both a master's degree in Automotive and Manufacturing and a Ph.D. in Mechanical Engineering from the University of Michigan, USA. After obtaining his bachelor's degree, Miguel led multiple projects developing manufacturing and assembly systems for alternative fuel vehicle programs. During his graduate studies at the University of Michigan, Miguel developed new methods for modeling and control of manufacturing systems for multi-objective optimization of plant floor operations. After graduation, Miguel joined General Motors Research and Development in June 2018 as a researcher. In his current role, Miguel has been able to capitalize on his strong technical and leadership skills to develop new technology in the field of robotics. His work aims to enable coordinated movement of multi-arm systems using artificial vision and force sensing data fusion.

Abstract: The automotive industry is constantly being challenged with increasing product variety, shorter life cycle, and demand uncertainty. In order to adapt in a highly competitive environment, the vehicle and components assembly plants need to have the flexibility to rapidly reconfigure and adapt to different products and production volumes. The concept of robotic assembly, where robots are used to place parts in the proper position was introduced as a solution to improve manufacturing flexibility while reducing cost and footprint. However, the use of robots for assembly presents some unique challenges particularly in perception and path planning that can affect the dimensional quality and throughput. Perception refers to the use of sensors such as cameras or laser radars to see and understand the part, process, and work environment conditions. The use of perception systems such as vision for robot guidance in precise positioning applications is often a challenge in a manufacturing environment due to inadequate lighting, poor part contrast, or limited field of view. Moreover, the vision system is expected to have high accuracy and reliability in order to maintain high levels of productivity. Some of the first developments of vision-based robotic assembly faced capability challenges mostly due to high cycle time and positioning errors. In the automotive industry the development of robotic assembly methods and control algorithms has focused largely on automotive body parts where 2D vision systems have been used to

locate part features and define the path of robot arms. Other perception alternatives such as 3D vision and a combination of 2D vision and laser readings have been introduced in various applications in order to improve accuracy and reduce cycle time. Moreover, the use of 2D vision might require additional robot movements that can be eliminated by using 3D vision, which can potentially help reduce cycle time. Recent developments in industrial robotics and artificial vision could help enable the next generation of robotic assembly systems. In this presentation a review of the challenges and opportunities of robotic assembly in the automotive industry is discussed. Also, examples of 2D and 3D vision for robotic assembly will be introduced. The focus will be to review the state-of-the-art of vision-based robot guidance for assembly and to highlight some key perception technology areas where research and development is required to enable robotic assembly of automotive body, powertrain, and battery assembly.

2.3.3 Technical Presentation 3

Title: Depth Quality Assessment at Close Range Using 3D Printed Fixtures

Author: Michele Pratusevich (Root AI)

Author Bio: Michele Pratusevich leads software and algorithm development as the Director of Software at Root AI, an agricultural robotics startup. Previously, Michele worked on computer vision, machine learning, and neural network applications targeted towards resource-starved systems at Amazon. At ICRA [International Conference on Robotics and Automation] 2019 Michele presented her work on close-range perception, showcasing a set of metrics for depth camera quality measurement and camera selection. She holds a Bachelor of Science (BS) and Masters of Engineering (Meng) in computer science and electrical engineering from the Massachusetts Institute of Technology (MIT).

Abstract: Mobile robots that manipulate their environments require high-accuracy scene understanding at close range. Typically, this understanding is achieved with RGBD [red, green, blue, and depth, or color + depth] cameras, but the evaluation process for selecting an appropriate RGBD camera for the application is minimally quantitative. Limited manufacturer-published metrics do not translate to observed quality in real-world cluttered environments, since quality is application-specific. To bridge the gap, we developed a method for quantitatively measuring depth quality using a set of extendable 3D printed fixtures that approximate real-world conditions. By framing depth quality as point cloud density and root mean square error (RMSE) from a known geometry, we present a method that is extendable by other system integrators for custom environments. We show a comparison of three cameras and present a case study for camera selection, provide reference meshes and analysis code, and discuss further extensions.

2.3.4 Technical Presentation 4

Title: Depth Camera Image Quality Definition and Measurement

Author: John Sweetser (Intel)

Author Bio: John Sweetser is currently a Computer Vision Engineer at Intel's RealSense CTO Group (previously known as Perceptual Computing). He has previously worked in various areas involving research and development, technology, and product development at start-ups (Templex Technology, ThinkOptics) and research labs (Sandia National Labs, Univ of Rochester) as well as Intel in a variety of areas involving Optical Engineering and Photonics. He has BS (Applied Physics) and MEng (EE) degrees from Cornell University and PhD from the University of Rochester's Institute of Optics.

Abstract: We will discuss the basic methods used at RealSense to evaluate the performance of depth cameras. This includes the definition of specific image quality metrics, methods, tools and test procedures for their measurement, typical performance standards, and examples of test results. Some discussion of qualitative image quality assessment as well as factors that can affect test results and overall performance will be included.

2.3.5 Technical Presentation 5

Title: High-resolution, high-speed 3D perception and sensing data streaming

Author: Song Zhang (Purdue University)

Author Bio: Dr. Song Zhang joined Purdue in January 2015 as an associate professor and was promoted to full professor in 2019. He received his Doctor of Philosophy (Ph.D.) degree in mechanical engineering from Stony Brook University in 2005. He is currently serving as the Assistant Head for Experiential Learning at the School of Mechanical Engineering, Purdue University. He received his Ph.D. degree in mechanical engineering from Stony Brook University in 2005; spent three years at Harvard as a postdoctoral fellow; and then worked at Iowa State University for 6 years before joining Purdue in January 2015. Dr. Zhang has over 200 publications. 15 of his journal articles were selected as cover page highlights. His publications have been cited over 8,900 citations with an h-index of 45. Besides being utilized in academia, the technologies developed by his team have been used by Radiohead (a rock band) to create a music video House of Cards; and by the law enforcement personnel to document crime scenes. He has received awards including AIAA [American Institute of Aeronautics and Astronautics] Best Paper Award, IEEE ROBIO [Institute of Electrical and Electronics Engineers International Conference on Robotics and Biomimetics] Best Conference Paper Award, Best of SIGGRAPH [Special Interest Group on Computer GRAPHics and Interactive Techniques] Disney Emerging Technologies Award, NSF CAREER [National Science Foundation Faculty Early Career Development Program] Award, Stony Brook University's "Forty under 40 Alumni Award", and CoE Early Career Faculty Research Excellence Award. He is currently serving as an associate editor for Optics and Lasers in Engineering, and as a technical editor for IEEE/ASME [Institute of Electrical and Electronics Engineers / American Society of Mechanical Engineers] Transactions on Mechatronics. He is a fellow of SPIE [International Society for Optics and Photonics] and OSA [The Optical Society].

Abstract: Advances in optical imaging and machine/computer vision have provided integrated smart sensing systems for intelligent systems; and advanced 3D perception techniques could have profound impact in the field of robotics. Our research addresses challenges in high-speed, high-resolution 3D perception and optical information processing. For example, we have developed a system that simultaneously captures, processes and displays 3D geometries at 30 Hz with over 300,000 measurement points per frame, which was unprecedented at that time (a decade ago). Our current research also explores novel means to stream/store enormously large 3D perception data by innovating geometry/video compression methods. The novel methods of converting 3D data to regular 2D counterparts offer the opportunity to leverage mature 2D data compression platform, achieving extremely high compression ratios without reinventing the whole data compression infrastructure. In this talk, I will present two platform technologies: 1) high-speed and high-resolution 3D perception; and 2) real-time 3D video compression and streaming. I will also cover some of the applications that we have been exploring including robotics, forensics, along with others.

2.3.6 Technical Presentation 6

Title: 3D Calibration and Perception for Robotic Scan-and-Plan Applications

Author: Joseph Schornak (Southwest Research Institute)

Author Bio: Joseph Schornak is a Research Engineer at Southwest Research Institute's Manufacturing and Robotics Technologies Department in San Antonio, TX and a contributor to the open-source ROS-Industrial metaproject. He has a Masters of Science (MS) in Robotics Engineering from Worcester Polytechnic Institute. His areas of interest include 3D perception, surface reconstruction, and robotic motion planning.

Abstract: Southwest Research Institute (SwRI) is a non-profit independent research and development institute located in San Antonio, TX. SwRI's Manufacturing and Robotics Technologies Department specializes in custom robotic solutions for advanced manufacturing applications. These systems rely on a wide variety of 3D sensors, including LIDAR [light detection and ranging], stereo cameras, time-of-flight cameras, and structured light scanners. Many of our ongoing challenges are centered around the calibration of these sensors, both intrinsically and in relation to the other sensors and robots that comprise each system. While we possess NIST-standard calibration artifacts, many of our calibration techniques and our methods of assessing the quality of data produced by each sensor began as ad-hoc solutions to implementation challenges encountered on specific systems, such as spatial error in 3D data and noise introduced by reflective surfaces. This talk will explore several case studies of perception-based robotic systems, as well as our current toolset for calibration and performance benchmarking.

2.3.7 Technical Presentation 7

Title: Using 3D vision to control robots in dirty, industrial environments

Author: Jared Glover (Capsen Robotics)

Author bio: Jared Glover is the Chief Executive Officer and co-founder of CapSen Robotics--a company that makes software to give robots more spatial intelligence. Jared received his Ph.D. in Computer Science from MIT in 2014, where he developed and applied new theoretical tools for processing 3D orientation information to applications in computer vision and robot manipulation. Prior to that, he completed his B.S. in Computer Science from Carnegie Mellon University, where he led a team developing robotic walkers for the Nursebot project. He has over 15 years of research experience in robotics and computer vision and over 400 paper citations. He is also a board member of Catalyst Connection, a private non-profit that provides consulting and training services to small manufacturers in southwestern Pennsylvania, and on advisory committees for the Advanced Robotics for Manufacturing (ARM) Institute, the Pittsburgh Robotics Network, and the NIST.

Abstract: CapSen Robotics writes 3D vision and motion planning software to give robots more spatial intelligence for manipulation tasks. The company's core product, CapSen PiC ("Pick in Clutter"), turns any industrial robot arm into a bin picking and machine tending cell. CapSen PiC handles parts of a wide range of sizes and shapes and can even disentangle picked objects. Our accompanying CapSen Scanner product captures 3D models in minutes, enabling the robot to quickly adapt to new jobs and parts. In this talk, I will discuss the practical challenges that robotics companies face in deploying 3D vision-guided robots in dirty, industrial settings. I will focus on two recent installations we've done. The first is in a wire & spring factory where our robot was tasked with picking metal hooks out of a bin, disentangling them (a first-of-its-kind capability in the robotics industry) and feeding them into a press. The second is for an application where novel parts must be scanned and then washed off. Both applications are in dirty environments and require the use of cutting-edge 3D vision algorithms. Yet they differ greatly in their requirements and methods. It is my hope that grounding our standards discussions with these practical case studies will help ensure that our metrics align with what end-users care most about--reliability!

2.4 Ranking Methodology

One of the goals of the workshop was to rank the proposed ideas for needed standards into a prioritized list, which would form the basis for the standards roadmap. The ranking methodology used was based on the American National Standards Institute (ANSI) Unmanned Aircraft Systems Standardization Collaborative's (UASSC) Standardization Roadmap for Unmanned Aircraft Systems (UAS) published in 2018 [1]. The method used to score and rank the ideas in Table 2 is the same method used in the Standardization Roadmap for Unmanned Aircraft Systems. The UASSC approach uses four critical elements (and a score from 1 to 3 for each) to rank each proposed idea for a standard (see Table 4).

2.4.1 Scoring

During the workshop, the participants were divided into three groups, and an attempt was made to maintain a mix of vendors and end users in each group. During the work sessions, each group was asked to score each idea from Table 2 based on the critical elements described in Table 4. The elements were given equal weight, and the total score for an idea was the sum of the scores for the four elements.

2.4.2 Ranking

The group ranking for an idea was determined as described in Table 4: High for scores between 10 and 12, Medium for scores between 7 and 9, and Low for scores between 4 and 6. In the group rankings, numerical values of 3, 2, or 1 were assigned to the High, Medium, and Low ideas, respectively. This allowed the three individual group rankings to be combined into a final ranking (from all three groups) for each idea.

Table 4. Prioritization Matrix adapted from [1].

Critical Element	Element Description	Scoring
Criticality	How important is the standard? How urgently is a standard or guidance needed? What would be the consequences if the standard were not completed or undertaken? A high score means the project is more critical.	3 = Critical 2 = Somewhat critical 1 = Not critical
Achievability (Time to Complete)	Does it make sense to develop this standard now, esp. when considered in relation to other standards? Is the standard already underway or is it a new standard? A high score means there's a good probability of completing the standard soon.	3 = Standard near completion 2 = Standard underway 1 = New standard
Scope (Investment in Resources)	Will the standard require a significant investment of time/work/money? Can it be completed with the information/tools/ resources currently available? Is pre-standardization research required? A high score means the standard can be completed without a significant additional investment of resources.	3 = Low resource requirement 2 = Med. resource requirement 1 = Resource intensive
Effect (Return on Investment)	What impact will the completed standard have on the industry? A high score means there are significant gains for the industry by completing the project.	3 = High return 2 = Medium return 1 = Low return
Score Rankings:	**High Priority** = a score of 10 to 12 **Medium Priority** = a score of 7 to 9 **Low Priority** = a score of 4 to 6	

2.5 Modified Ranking

The equal weighting of the four elements in the UASSC methodology was questioned by the workshop participants who felt that the Criticality and Effect elements were of higher importance than the Achievability and Scope elements. Therefore, the participants proposed and compared three different methods of determining the final ranking for each idea from the individual group rankings and scores. A description of each method follows.

- **Method 1**: The final ranking for each idea was the average of the group rankings based on using all four elements with equal weights. The final ranking was then a real number ranging from 1 (Low) to 3 (High).

- **Method 2**: The final ranking for each idea was based on the average of the scores from all three groups. This method is similar to Method 1 in that it is based on all four elements with equal weight. However, unlike Method 1, which uses the average of the three group rankings, the second method uses the average of the *scores* from each of the three groups to determine the final rankings instead. Therefore, if the average score from all three groups was between 10 and 12, the final ranking was set to 3 (High), if it was between 7 and 9, the final ranking was set to 2 (Medium), and if it was between 4 and 6, the final ranking was set to 1 (Low). In this method, the final ranking for each idea is either 1, 2, or 3.

- **Method 3**: The final ranking for each idea was based solely on the scores from two of the four elements: Criticality and Effect. For each group, and for each idea, only the scores for the Criticality and Effect were summed. Then, the average of the scores from the three groups was used as the final ranking, and the final rankings ranged from 2 (Low) to 6 (High).

3 Work Sessions

A total of three work sessions of approximately two hours each were held during the workshop. Descriptions of these sessions are presented below.

3.1 Work Session 1: Prioritizing the 27 Ideas

During the first work session on Day 1 of the workshop, participants were divided into three groups as described in Section 2.4.1. Each group was tasked with scoring the 27 original ideas for standards that are needed for 3D perception systems that were developed prior to the workshop. The scoring was based on the critical elements described in Table 4. The groups were given 5 minutes to score each idea, which was first described to the attendees by the moderator.

3.2 Work Session 2: Finish Prioritizing the 27 Ideas and Develop and Prioritize New Ideas

During the second work session on Day 1, workshop participants remained in the same groups as assigned in the Work Session 1. Each of the three groups were given 30 minutes to finalize the scoring of the original 27 ideas. The groups were then given another 30 minutes to come up with any new ideas that were not represented in the original list of 27 ideas. Finally, the new ideas generated by each group were consolidated into a single list and all three groups were then given 15 minutes to score the new ideas (including the 3 new ideas listed in Table 2).

3.3 Work Session 3: Develop the Top 9 Ideas into ASTM Work Items

The third, and final work session took place on Day 2 of the workshop. During this work session the first ranking method (Method 1) presented in Section 2.4 was described by the moderator to the participants. The intent was to then select the top 9 ideas from the resulting list of ranked ideas and to develop those ideas further.

However, based on feedback from the participants, the Method 2 and Method 3 rankings were calculated by the workshop organizers and presented to the participants. The participants then debated which ranking method to use to select the 9 highest priority ideas and finally settled on a hybrid approach that is described in Section 4.2.

3.4 Group Scores and Rankings

The individual group scores and ranking may be found in Appendix C.

4 Workshop Results

4.1 Ranked Results

The rankings for the original ideas (27) and the new ideas (12) for each group at the workshop are shown in Table 5 and

Table 6 using two different sorting methods. In Table 5, the ranked ideas were sorted from high to low based on Method 1 ranking. In

Table 6, the ranked ideas were sorted from high to low based on Method 3 ranking. Since the sorted ranking did not change very much between Methods 1 and 2, sorting based on Method 2 is not shown.

In Table 7**Error! Reference source not found.**, the top 10 ranked ideas are listed for the three methods. As seen in Table 7**Error! Reference source not found.**, the top 10 ideas in Method 1 and Method 2 are the same (but differ slightly in priority), whereas the ideas were quite different in Method 3.

Table 5: Ideas ranked via the 3 methods and sorted based on the ranking in Method 1

Ideas	RANKING via Method 1 (3 = High, 1 = Low)	RANKING via Method 2 (3 = High, 1 = Low)	RANKING via Method 3 (6 = High, 3 = Low)
Eye safety over FOV	3.00	3	5.33
Changes in performance throughout a perception system's field-of-view (FOV)	2.67	2	6.00
Measurement volume specification/verification	2.67	2	6.00
Standard reference objects or artifacts	2.67	3	5.67
Error against traceable targets	2.67	2	5.67
Interoperability	2.67	2	5.33
Functional safety	2.67	2	5.00
Ambient conditions	2.33	2	6.00
Pointcloud XYZ resolution	2.33	2	6.00
2D image XY resolution	2.33	2	6.00
Part position resolution	2.33	2	6.00
Part orientation resolution	2.33	2	6.00
Reliability & Robustness	2.33	2	6.00
Latency	2.33	2	5.00
Depth map XYZ resolution	2.00	2	6.00
Output quality	2.00	2	6.00
Ability to resolve geometric features	2.00	2	6.00
Standard robot platform for complete system testing	2.00	2	6.00
Repeatability	2.00	2	5.33
Cycle time	2.00	2	5.00
Performance due to part material properties	2.00	2	5.00
Performance due to part surface properties	2.00	2	5.00
Data compression	2.00	2	5.00
Calibration quality	2.00	2	5.00
Static performance	2.00	2	4.67
Dynamic performance	2.00	2	4.33
Power connector interface	2.00	2	4.00
Time synchronization	2.00	2	3.67
Power requirements	2.00	2	3.67

Ideas	RANKING via Method 1 (3 = High, 1 = Low)	RANKING via Method 2 (3 = High, 1 = Low)	RANKING via Method 3 (6 = High, 3 = Low)
Frame rate	2.00	2	3.67
Depth error	1.67	2	5.00
Part reflectance	1.67	2	4.33
Computation power of host computer	1.67	2	4.33
XYZ linearity	1.67	1	3.33
Performance due to cluttered versus uncluttered scenes	1.33	1	4.33
Performance due to occlusions	1.33	1	4.00
System-to-part suitability	1.33	1	3.00
Temperature stability	1.33	1	2.67
Bit precision resolution	1.33	1	2.33

Table 6: Ideas ranked via the 3 methods and sorted based on the scores in Method 3

Ideas	RANKING via Method 1 (3 = High, 1 = Low)	RANKING via Method 2 (3 = High, 1 = Low)	RANKING via Method 3 (6 = High, 3 = Low)
Changes in performance throughout a perception system's field-of-view (FOV)	2.67	2	6.00
Measurement volume specification/verification	2.67	2	6.00
Ambient conditions	2.33	2	6.00
Pointcloud XYZ resolution	2.33	2	6.00
2D image XY resolution	2.33	2	6.00
Part position resolution	2.33	2	6.00
Part orientation resolution	2.33	2	6.00
Reliability & Robustness	2.33	2	6.00
Depth map XYZ resolution	2.00	2	6.00
Output quality	2.00	2	6.00
Ability to resolve geometric features	2.00	2	6.00
Standard robot platform for complete system testing	2.00	2	6.00
Standard reference objects or artifacts	2.67	3	5.67
Error against traceable targets	2.67	2	5.67
Eye safety over FOV	3.00	3	5.33
Interoperability	2.67	2	5.33
Repeatability	2.00	2	5.33
Functional safety	2.67	2	5.00
Latency	2.33	2	5.00
Cycle time	2.00	2	5.00
Performance due to part material properties	2.00	2	5.00
Performance due to part surface properties	2.00	2	5.00
Data compression	2.00	2	5.00
Calibration quality	2.00	2	5.00
Depth error	1.67	2	5.00
Static performance	2.00	2	4.67
Dynamic performance	2.00	2	4.33
Part reflectance	1.67	2	4.33
Computation power of host computer	1.67	2	4.33
Performance due to cluttered versus uncluttered scenes	1.33	1	4.33

Ideas	RANKING via Method 1 (3 = High, 1 = Low)	RANKING via Method 2 (3 = High, 1 = Low)	RANKING via Method 3 (6 = High, 3 = Low)
Power connector interface	2.00	2	4.00
Performance due to occlusions	1.33	1	4.00
Time synchronization	2.00	2	3.67
Power requirements	2.00	2	3.67
Frame rate	2.00	2	3.67
XYZ linearity	1.67	1	3.33
System-to-part suitability	1.33	1	3.00
Temperature stability	1.33	1	2.67
Bit precision resolution	1.33	1	2.33

Table 7: Top 10 ranked ideas

	Method 1	**Method 2**	**Method 3**
1	Eye safety over FOV	Eye safety over FOV	Changes in performance throughout a perception system's field-of-view (FOV)
2	Changes in performance throughout a perception system's field-of-view (FOV)	Standard reference objects or artifacts	Measurement volume specification/verification
3	Measurement volume specification/verification	Changes in performance throughout a perception system's field-of-view (FOV)	Ambient conditions
4	Standard reference objects or artifacts	Measurement volume specification/verification	Pointcloud XYZ resolution
5	Error against traceable targets	Error against traceable targets	2D image XY resolution
6	Interoperability	Interoperability	Part position resolution
7	Functional safety	Functional safety	Part orientation resolution
8	Ambient conditions	Ambient conditions	Reliability & Robustness
9	Pointcloud XYZ resolution	Pointcloud XYZ resolution	Depth map XYZ resolution
10	2D image XY resolution	2D image XY resolution	Output quality

4.2 Work Items

A desired outcome of the workshop was to extend the top ideas into work items[6] so that they could be developed into standards. It was anticipated that task groups would be formed using these work items as starting points, and these task groups would work towards fully developing them into standards. Each standard could be identified as a specification, test method, practice, guide, classification, or terminology.

To aid these future task groups, participants were asked to come up with the following information for each work item:

- Title
- Proposed scope
    - For a Test Method, ASTM describes the scope as follows [2]:
        - Include information relating to the purpose of the test method. State if the method is quantitative or qualitative, and any known limitations. Concisely state the property or constituent that is being determined and the materials that can be analyzed. Include, where applicable, the analytical technique, for example, gas chromatography, and whether the test is performed in the laboratory, field, or on-line.
    - For a Specification ASTM describes the scope as follows [2]:
        - Include information relating to the purpose of the specification. Concisely state the materials, products, systems, or services to which the specification applies and any known limitations. Include, where applicable, the intended use of the specification. Do not include references to trademarks.
    - For Practices and Guides ASTM describes the scope as follows [2]:
        - Include information relating to the purpose of the practice or guide and to what it applies. Clearly state any limitations of the practice or guide.
- Rationale (explains why the standard is needed, how it will be used, and who the users will be)
- Technical Contact
- Additional contributors
- Target date for Subcommittee or Concurrent Ballot
- Type of standard:
    - Specification
    - Test method
    - Practice/Guide
    - Classification
    - Terminology

[6] "A Work Item (WK) may be a new standard or a revision to an existing standard that is under development by a committee." (From https://www.astm.org/DATABASE.CART/whatisaworkitem.html)

- Keywords
- List of existing standards and why it is necessary to develop an ASTM standard
- List of other ASTM committees or key outside organizations that should be informed of the activity

Since the following three ideas were among the top ten ideas in all three ranking methods, they were slated as work items 1 to 3:

1. Ambient Conditions
2. Changes in Performance throughout a Perception System's FOV
3. Measurement Volume Specification/Verification

For work items 4 to 9, the meeting participants were asked to pick six of the remaining ideas from **Error! Reference source not found.**. The participants selected five ideas from Table 7 and decided that "Latency" (although it was not part of the top 10 ranked ideas in any of the three ranking methods) was sufficiently important to include as part of the nine work items. Therefore, the remaining six work items were:

4. Standard reference objects or artifacts
5. Latency
6. XYZ resolution (for depth maps and pointclouds),
7. Part position resolution
8. Part orientation resolution
9. Output quality

Due to time limitations, only six of the nine work items described above were developed further and none of the groups were able to supply information for Target date, List of existing standards, and List of other ASTM committees or key outside organizations, and therefore, these rows are left out of the six work items listed in Table 8 to

Table 13.

Table 8: Work Item 1 - Ambient Conditions

Title	Practice for considering the effects of lighting on Output of a 3D Perception System
What is the Proposed Scope? (See Scope worksheet)	Apply to all systems. For instruments with ranges up to 3 m. Frequency change of the intensity of the lighting, spectral distribution of the light. Constant conditions, within frame (one data acquisition), across frames, high frequency change (e.g. flash), low frequency (gradual change in lighting). This standard determines the variation in the performance of a system under various lighting conditions. This standard will not be defining any metrics as these metrics will be defined in the standard developed for a particular performance characteristic.
Rationale for New Standard (explain why the standard is needed, how it will be used and who the users are)	Main cause of failures of perception systems is due to lighting. Users: End users who want to use the systems in varying lighting conditions, manufacturers can use it to improve their sensor.
Who will be the Technical Contact for this Work Item?	
Who will be the other contributors to this Work Item? (Name, affiliation, email address)	Benjamin Carrier (NRC), Michele Pratusevich, Jared Glover, Gil Summy, Yoshi Ohno (technical consultant when needed), Marc-Antoine Drouin, NIST will provide support
What is the type of Standard?	
Please supply useful Keywords not in the Scope that users would employ to search for this Work Item	

Table 9: Work Item 2 - Changes in Performance throughout a Perception System's FOV

Title	Test method for measuring the Performance of a 3D perception system across the specified FOV
What is the Proposed Scope? (See Scope worksheet)	A quantitative test method for evaluating the 3D measurement performance across a specified volume of 3D perception systems. Where performance includes items such as fill ratio, spatial density, spatial noise, temporal noise, z-accuracy, spatial resolution, and minimum detectable object size.
Rationale for New Standard (explain why the standard is needed, how it will be used and who the users are)	Existing standards do not fully define the common performance definitions and characteristics of 3D perception systems
Who will be the Technical Contact for this Work Item? (Name, affiliation, email address)	John Sweetser
Who will be the other contributors to this Work Item? (Name, affiliation, email address)	John Horst, Remus Boca, Jared Glover, Miguel Saez, Peter Walecki, Etienne Del Torchio, Brent Fisher, John Sweetser, Prem Rachakonda
What is the type of Standard?	
Please supply useful Keywords (separated by commas) not in the Scope that users would employ to search for this Work Item	

Table 10: Work Item 3 - Measurement Volume Specification/Verification

Title	Test method for verification of a 3D perception system's working performance volume
What is the Proposed Scope? (See Scope worksheet)	The standard escribes a quantitative test method for verifying the working performance volume of a 3D perception system of the specified range. The term "working performance volume" refers the region within the system's satisfied minimum performance threshold. This test method only applies to 3D perception systems that has specifications of working volume.
Rationale for New Standard (explain why the standard is needed, how it will be used and who the users are)	
Who will be the Technical Contact for this Work Item?	Felix Thouin
Who will be the other contributors to this Work Item? (Name, affiliation, email address)	Brent Fisher, Leung Shiu, Helen Qiao
What is the type of Standard?	
Please supply useful Keywords (separated by commas) not in the Scope that users would employ to search for this Work Item	How to define working performance volume (e.g. in which coordinate frame in x, y, or spherical coordinate), key features to describe the working volume (e.g. standoff distance, maximum range in space), how to define a set of performance metrics that need threshold (maybe based on user's application), how to define the way to test the selected performance.

Table 11: Work Item 4 - Standard Reference Objects or Artifacts

Title	Develop standard reference artifact(s) to evaluate the performance of a 3D Perception System
What is the Proposed Scope? (See Scope worksheet)	Specify artifacts with varying materials, color, texture, reflectivity, size, geometry, features to evaluate the performance of a 3D Perception System. This will not include artifacts for the resolution test.
Rationale for New Standard (explain why the standard is needed, how it will be used and who the users are)	Provide a standard artifact(s) to evaluate performance of perception systems.
Who will be the Technical Contact for this Work Item?	
Who will be the other contributors to this Work Item? (Name, affiliation, email address)	Remus Boca, Michele P., Miguel Saez, Gil Summy, Marc-Antoine Drouin, Benjamin Carrier, Prem Rachakonda, with NIST support
What is the type of Standard?	
Please supply useful Keywords (separated by commas) not in the Scope that users would employ to search for this Work Item	

Table 12: Work Item 5 - Latency

Title	Terminology for describing the time delay in 3D sensor output
What is the Proposed Scope? (See Scope worksheet)	Standards for defining the latency and frame rate of a 3D sensor, beginning with the "start time," i.e., the moment of time that some sensor component (e.g., and image chip) begins acquiring sensor data, and ending with the availability of useable 3D output
Rationale for New Standard (explain why the standard is needed, how it will be used and who the users are)	A wide variety of terminology is used, and vendors and manufacturers measure the timing quantities differently.
Who will be the Technical Contact for this Work Item?	Jared Glover
Who will be the other contributors to this Work Item? (Name, affiliation, email address)	John Sweeter, Etienne Del Torchio, John Horst, Prem Rachakonda
What is the type of Standard?	Terminology for ??
Please supply useful Keywords (separated by commas) not in the Scope that users would employ to search for this Work Item	

Table 13: Work Item 6 - XYZ Resolution (depth map, pointcloud)

Title	Test methods for determination of a 3D perception system's point wise spatial resolution
What is the Proposed Scope? (See Scope worksheet)	This test method covers the performance evaluation of 3D perception system's point wise spatial resolution. The term "point wise spatial resolution" refers the minimum distinguishable distance between two points within a specified volume.
Rationale for New Standard (explain why the standard is needed, how it will be used and who the users are)	Define what's the point wise spatial resolution. define an artifact (what shape to use) to test small steps in x, y, and z direction. test method: place the target at different pose and distance. question of a small section of the sensor or a big area of the sensor
Who will be the Technical Contact for this Work Item? (Name, affiliation, email address)	Felix Thouin, Joseph Schornak
Who will be the other contributors to this Work Item? (Name, affiliation, email address)	Joseph Schornak, Peter Walecki, Song Zhang, Brent Fisher, Leung Shiu, Helen Qiao
What is the type of Standard?	
Please supply useful Keywords (separated by commas) not in the Scope that users would employ to search for this Work Item	Define what's the point wise spatial resolution, define an artifact (shape e.g. flat surface, sphere, etc. to use) to test small steps in x, y, and z direction, define test method to test the artifact at different pose and distance, define the use of a small section of the sensor or a big area of the sensor.

5 Panel Discussions

The panel discussion session allowed participants to bring up topics of interest to them and included discussions on the current state of robotic perception technology related to applications in manufacturing environments. Panelists provided insights on perception technology and systems, in particular current capabilities, best practices, and challenges, as well as technology gaps and limitations. The moderator and panelists are listed. Key points are summarized below.

5.1 Panelists for Robotic Perception Capabilities, Best Practices, Challenges, and Needs

The moderator for the Panel Discussion session was Dr. Kamel Saidi from NIST. The panelists were:

- John Sweetser, Intel Corp.
- Jared Glover, CapSen Robotics
- Miguel Saez, General Motors
- Remus Boca, ABB
- Joseph Schornak, SwRI
- Song Zhang, Purdue

5.2 Support for Industrial Applications

A question was raised about vendor support in industrial applications. For example, the National Research Council (NRC, Canada) currently has a lot of industrial applications. If NRC used the Intel RealSense D400 cameras for their applications, would these cameras stand up to the physical demands of such an application?

Response: Intel is developing new camera models with longer range, Ingress Protection Ratings of 65 (IP65), and more robustness to dust, temperature, and vibration to support industrial applications.

5.3 Standard Development Time Frame

Questions were raised about the time frame for the prioritized list and about whether five years would be a proper time frame for the high priority standards.

Responses: Five years seems to be a long time for high priority. Technologies are still evolving. Fast updates are needed to catch up. Terminology and specification should be a shorter time frame than test methods.

5.4 System Specifications

A question was raised about how to present data in the specification sheet until standards are available.

Responses: Examples include having graphical charts/video to show users the specifications. Aerospace manufacturers often have internal metrology departments for performing verification tests and giving certificates. There are probably some good lessons to be learned from this industry.

5.5 Who should be Involved in the Development of the Standards?

Questions were raised about who should be involved in developing the standards and about what it would take to get vendors to agree with the terminology.

Responses: Vendors, integrators, and end-users should be involved in this standard. Involving end-users is important. For example, if Caterpillar, John Deere, and Steel Case were to ask for standards, that would push vendors' efforts in standards development. Another example is Mass Robotics – a startup incubator in Massachusetts, who may have interest in testing different solutions to support the standards development. The U.S. Army Corp of Engineers could be another candidate too. Also, showcases at end-user site could bring more attention from integrators and vendors.

5.6 Channels to Advertising the Standard Development Work

Questions: Before we publish the workshop report, who should be involved in the review? How should we "advertise" the next perception workshop?

Responses: AIA/RIA (Association for Advancing Automation / Robotics Industries Association), members of the ROS (Robot Operating System) Industrial consortium, integrators, and vendors should be exposed to the review. Channels to advertising the next perception workshop include publishing an article (e.g. Quality Magazine), with a report of this workshop; a booth at the Vision 2020 Show; social media; getting vendors to have some common messaging in their booths about the standards; holding a workshop concurrent with conferences/shows (e.g. ASME MSEC). For example, ARM (Advanced Robotics for Manufacturing) Institute of Pittsburgh could host a workshop.

5.7 Getting more Involvement from Academia

Questions were raised about how to get more academia involved and about whether competitions would help.

Responses: Researchers need to know NIST is working in the related research areas. Researchers are happy to give input, or use artifacts developed from the standards. Academia has difficulty developing artifacts by themselves. Competitions could be beneficial, but it would depend on the circumstance as it could be very costly to a university. Competitions work for undergraduate students with less effort. Graduate students need to find financial support to work on the competition and competitions with financial incentives would be helpful.

5.8 Standards Development Priority

A question was asked about which standards we would pick if we could only work on one or two of the six standards.

Responses: Developing a standard physical artifact got the most votes. It is important to develop artifacts to understand the accuracy of the system. Even if NIST develops an artifact without consensus, it may start to get people's interest. The artifact can be modified later to accommodate other needs. Developing metrics and test methods for evaluating the accuracy of an instrument throughout its FOV received the second most number of votes.

5.9 Emerging Technologies that Impact Future Perception Applications

A question was raised about what the important emerging technologies are that will impact future perception applications.

Panelists expressed that some of the desirable advances in sensor technology include adaptive autoexposure, auto zoom, higher dynamic range, and multiple focus technology. It was pointed out that autonomous vehicles are driving innovations (e.g., Mobile Eye) and that having a large number of users will drive the development of new technologies.

Terahertz imaging is an emerging and significant nondestructive evaluation technique used for dielectric (nonconducting - i.e., insulator) materials analysis and quality control that could be used to see inside an object. Intel is developing a scanning LIDAR that works at a long-range and that comes at a low cost. Subwavelength imaging provides the ability to see details of an object or organism below the wavelength of visible light, to have the capability to observe, in real-time, below 200 nm.

5.10 Artificial Intelligence (AI) and Perception Technology

A question was raised about how artificial intelligence (AI) is changing (or will change) available perception system capabilities.

Responses: AI is still in the research stage. It lacks transparency from the vendor side. It is more like embedded algorithms on chips, using deep learning. Another option is that vendors don't use deep learning but use a more simplified machine learning using a small dataset instead of a very large dataset. Datasets for stereo vision systems specifically for manufacturing scenes and objects are needed.

5.11 Action Items

- Set up the six work items under ASTM E57 and form task groups to work on each one.
- Publish a report on the workshop.
- Give a presentation at the Vision Show about the workshop.
- Organize a meeting at the Vision Show in June in Boston.
- Publicize the work of the ASTM AC475 working group through the following venues:
    - Publish an article in Quality magazine about the workshop
    - ASME Manufacturing Science and Engineering Conference (MSEC) meeting in June in Cincinnati.
    - Posts on social media.
    - Develop a common message about ASTM E57 3D imaging standards that can be used at vendor booths and other literature.
    - Develop and give free webinars.
- Develop and send the standards roadmap to different organizations for feedback.

6 References

[1] ANSI Standardization Roadmap for Unmanned Aircraft Systems, https://www.ansi.org/news_publications/news_story?menuid=7&articleid=58757077-aeb7-4554-b359-4aa34ae8881d (2018)

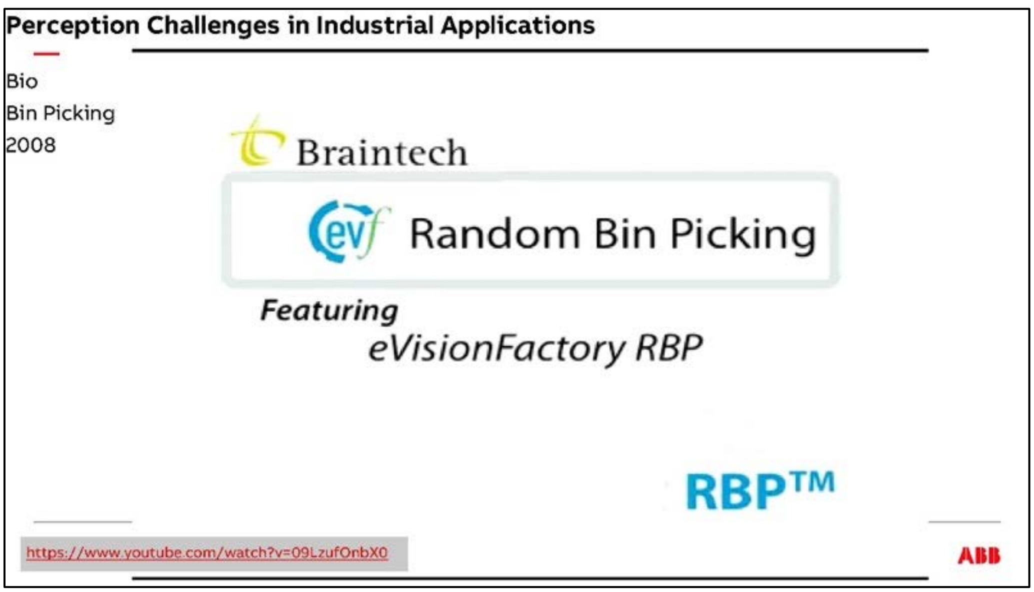
[2] Form and Style for ASTM Standards, ASTM International, West Conshohocken, PA, 2018, https://www.astm.org/FormStyle_for_ASTM_STDS.html

Appendix A: List of Participants

Antonishek, Brian	NIST	brian.antonishek@nist.gov
Bagchi, Shelly	NIST	shelly.bagchi@nist.gov
Billotte, Bill	ASTM	wbillotte@astm.org
Boca, Remus	ABB Inc	remus.boca@us.abb.com
Cheok, Gerry	NIST	cheok@nist.gov
Chu, Wei	Self-Employed	larrych88@gmail.com
del Torchio, Etienne	Visio Nerf	edeltorchio@visionerf.com
Drouin, Marc-Antoine	National Research Council Canada	marc-antoine.drouin@nrc-cnrc.gc.ca
Fisher, Brent	Sense Photonics	brentrf@gmail.com
Glover, Jared	Capsen Robotics	jared.glover@capsenrobotics.com
Horst, John	NIST	john.horst@nist.gov
Kimble, Kenneth	NIST	kenneth.kimble@nist.gov
Li, Xiang	NIST	xiang.n.li@nist.gov
Messina, Elena	NIST	elena.messina@nist.gov
Pratusevich, Michele	Root AI	mprat@root-ai.com
Qiao, Helen	NIST	guixiu.qiao@nist.gov
Rachakonda, Prem	NIST	prem.rachakonda@nist.gov
Saez, Miguel	General Motors	miguel.saez@gm.com
Saidi, Kamel	NIST	kamel.saidi@nist.gov
Schornak, Joseph	Southwest Research Institute	joseph.schornak@swri.org
Shiu, Leung	USPS	lshiu@usps.gov
Summy, Gil	Airy3D	gil.summy@airy3d.com
Sweetser, John	Intel	john.n.sweetser@intel.com
Thouin, Felix	Airy:3D	felix.thouin@airy3d.com
Walecki, Peter	Brown University Engineering Dept.	peter_walecki@brown.edu
Xiang, Dan	X-wave Innovations, Inc.	dxiang@x-waveinnovations.com
Yang, Bingcheng	May Solutions LLC	ayang@MaySolutionLLC.com
Zhang, Song	Purdue University	szhang15@purdue.edu

Appendix B: Presentations

B.1. Perception Challenges for Industrial Applications by Remus Boca, ABB



Perception Challenges in Industrial Applications

Agenda

As the world moves towards autonomy, the sensing and perception are becoming (critical) important.

Industrial applications have their own challenges as they operate in possible harsh environments, they require continuous and robust operation, need to accommodate unstructured environments, acknowledge a wide range of states and unexpected events.

This talk presents perception needs and challenges across many industries such as ports, mining, transformer inspection and robotics.

Perception Challenges in Industrial Applications

Ports

Multiple vision modalities for different tasks
- **remote control**
- **pick / drop container**
- **safety**

Outdoor operation
- **Enclosure is needed**
- **environment conditions : day / night; sunny, cloudy, rain, snow, wind; humidity**

Working range
- **from 1m to 30m (stacking crane)**
- **from 1m to 85 m (ship to shore)**

Accuracy medium to low (cm)

Access to sensors is required

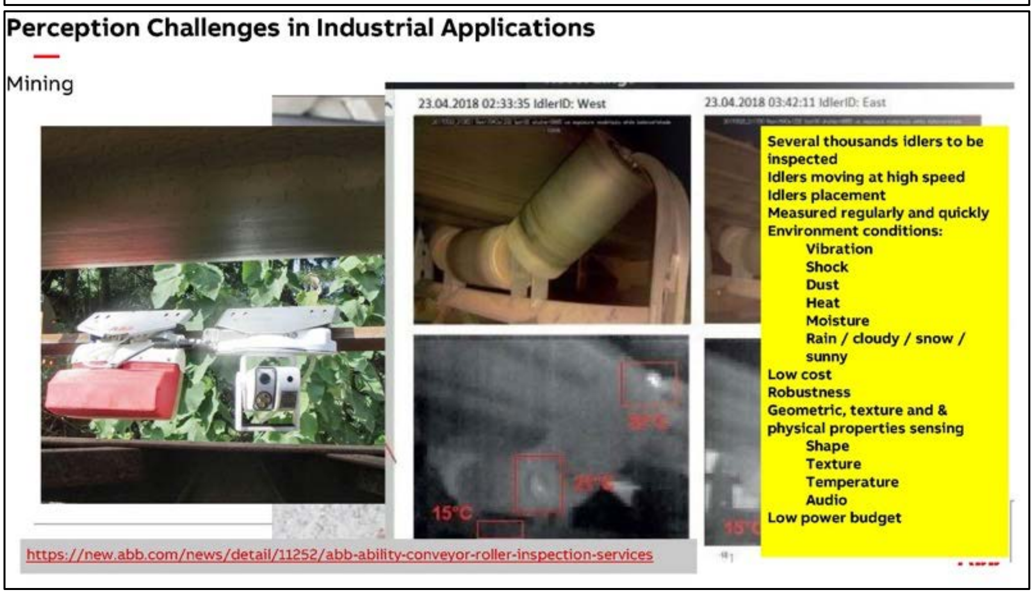
Objects of interest wear and tear
- **container & container stacks**
- **chassis**
- **ships / bays**

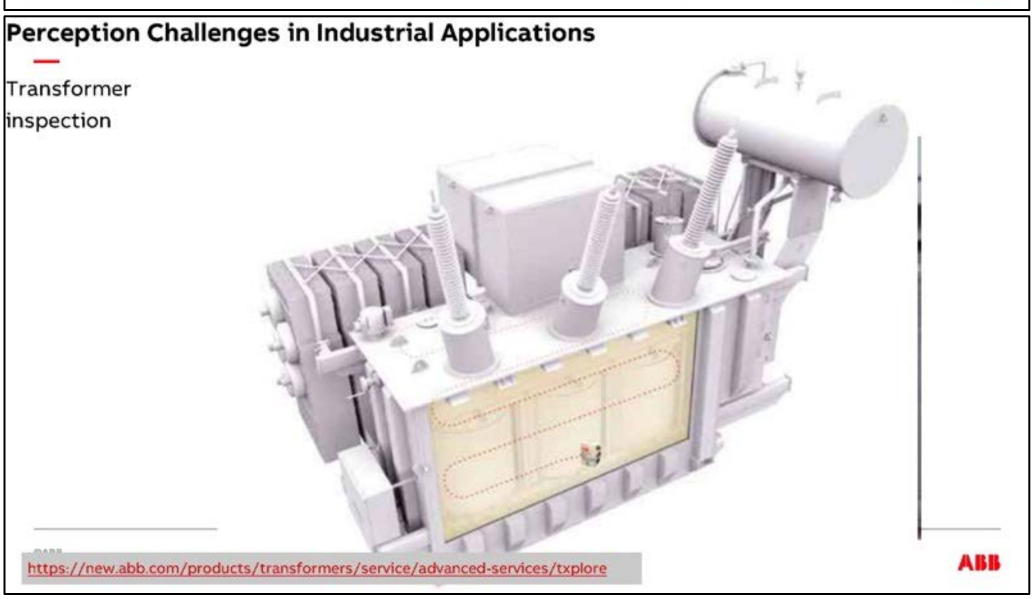
Perception Challenges in Industrial Applications

Ports

https://new.abb.com/ports


Perception Challenges in Industrial Applications
Mining
https://new.abb.com/news/detail/11252/abb-ability-conveyor-roller-inspection-services
ABB


Perception Challenges in Industrial Applications
Mining
23.04.2018 02:33:35 IdlerID: West
23.04.2018 03:42:11 IdlerID: East
Several thousands idlers to be inspected
Idlers moving at high speed
Idlers placement
Measured regularly and quickly
Environment conditions:
Vibration
Shock
Dust
Heat
Moisture
Rain / cloudy / snow / sunny
Low cost
Robustness
Geometric, texture and & physical properties sensing
Shape
Texture
Temperature
Audio
Low power budget
15°C
https://new.abb.com/news/detail/11252/abb-ability-conveyor-roller-inspection-services


Perception Challenges in Industrial Applications
Transformer inspection
https://new.abb.com/products/transformers/service/advanced-services/txplore
ABB

Perception Challenges in Industrial Applications

Transformer inspection

- Operating environment
  - It can be contaminated
  - Decontaminate after every use
  - Wide range of temperatures
- Resolution for quality images
- Field of view / surround
- Short distance
- Lighting
- More than one sensor
- Many types of defects
  - Appearance
  - Geometric
- Wireless communication
- Sensors are critical to operate while submersed (expensive to recover the device)
  - redundant cameras

https://new.abb.com/products/transformers/service/advanced-services/txplore

ABB

Perception Challenges in Industrial Applications

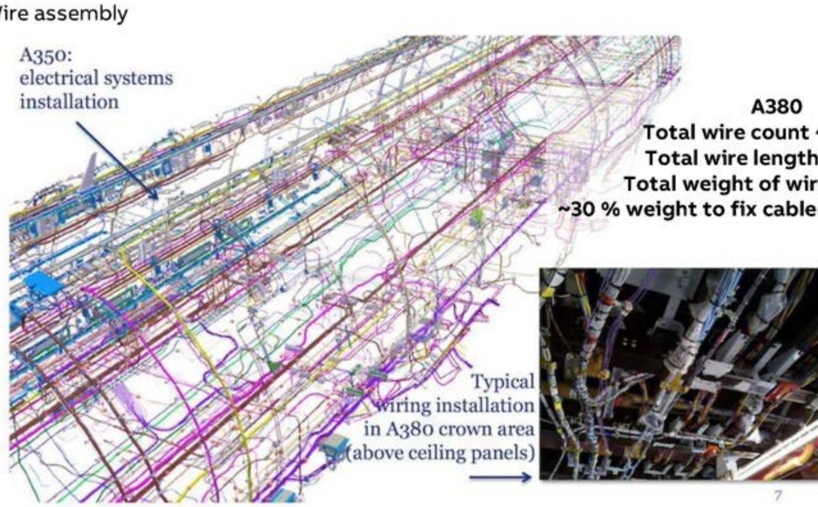
Transformer Inspection
ABB Ability TXplore

https://www.youtube.com/watch?v=OclzoEc8plE

ABB

Perception Challenges in Industrial Applications

Wire assembly



A380
Total wire count ~100 000
Total wire length ~470Km
Total weight of wires 5700 Kg
~30 % weight to fix cables to the structure

7

ABB

Perception Challenges in Industrial Applications

Wire assembly

Luxury car
Total wire count ~1500
Total wire length ~1 Mile
Total weight of wires 150 #

- Accuracy sub-mm
- High resolution
- High frame rate
- Small and large field of view
- Small footprint
- Short range
- Quality components (lens, global shutter)
- Multiple sensors working on the same field of view
- Low cost
- Lifecycle as long as a robot
- Easy diagnostic
- Industrial communication protocols
- Easy customization

©ABB
December 3, 2019 | Slide 12

ABB

Perception Challenges in Industrial Applications

Ports

Multiple modalities for different tasks
- remote control
- pick / drop container
- safety

Outdoor operation
- enclosure
- environment conditions : day / night; sunny, cloudy, rain, snow, wind; humidity

Working range
- from 1m to 30m (stacking crane)
- from 1m to 85 m (ship to shore)

Accuracy medium to low (cm)
Access to sensors
Objects of interest wear and tear
- container & container stacks
- chassis
- ships / bays

Mining

Several thousands idlers to be inspected
Idlers moving at high speed
Idlers placement
Measured regularly and quickly
Outdoor & environment conditions:
- Vibration
- Shock
- Dust
- Heat
- Moisture
- Rain / cloudy / snow / sunny

Low cost
Robustness
Geometric, texture & physical properties sensing
- Shape
- Texture
- Temperature
- Audio

Low power budget

Transformer inspection

Operating environment
- It can be contaminated
- Decontaminate after every use
- Wide range of temperatures

Resolution for quality images
Field of view / surround
Short distance
Lighting
More than one sensors
Many types of defects
- Appearance
- Geometric

Wireless communication
Sensors are critical to operate while submersed (expensive to recover the device)
redundant cameras

Wire assembly

- Accuracy sub-mm
- High resolution
- High frame rate
- Small and large field of view
- Small footprint
- Short range
- Quality components (lens, global shutter)
- Multiple sensors working on the same field of view
- Low cost
- Lifecycle as long as a robot
- Easy diagnostic
- Industrial communication protocols
- Easy customization

©ABB
December 3, 2019 | Slide 13

ABB

B.2. Robotic Assembly: Challenges and Opportunities in the Automotive Industry by Miguel Saez, General Motors

ROBOTIC ASSEMBLY: CHALLENGES AND OPPORTUNITIES

Miguel Saez Ph.D.
Researcher
Manufacturing Systems Research Lab

RESEARCH & DEVELOPMENT

KEY TECHNOLOGY FEATURES

- Perception: Use of sensors to understand the work environment and process requirements:
  - Exteroceptive:
    - Digital image: One or multiple cameras to identify target features
    - Laser: Point or line readings to create a point cloud or line
  - Enteroreceptive:
    - Load cells: Measure forces and/or torques in robot arm or end effector
    - Motor torque: Motor current feedback
- Robot coordination
  - Synchronous motions: Share positional and target information to execute separate tasks
  - Asynchronous motions: Share positional and target information to execute the same task
- Robot accuracy: Identify the actual and target pose of the robot arm

RESEARCH & DEVELOPMENT

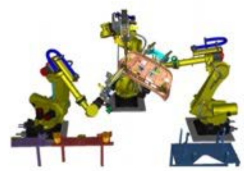
15

INTRODUCTION

What is Robotic Assembly?

Academia:

- Fixtureless Assembly: Use of robots to place parts in the proper position without the need of a dedicated fixture (Hoska, 1988)
- Vision-guided positioning : Interaction of multiple robots to position and hold parts to perform a task (Bone & Capson, 2003) (Novakovic et al. 2017)
- Coordinated motion: Control joint position and torque of multiple robots holding a single parts (Gueaieb et al., 2007) (Uchiyama et al, 1987)



Industry:

- Clutch insertion using vision and force sensing (Gravel et al., 2008)
- Robot-to-Robot handoff at GM body shop plants
- "Open-loop" positioning of sheet metal parts using multiple robot arms

GM RESEARCH & DEVELOPMENT

3

ASSEMBLY OPERATIONS

- Body Assembly:

- Vehicle Assembly

- Powertrain assembly

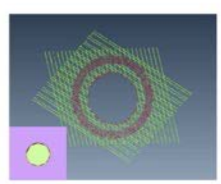
- Battery Assembly

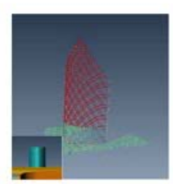
GM RESEARCH & DEVELOPMENT

7

PERCEPTION IN MANUFACTURING

2D:
- Digital image: Detection of objects positioned in a work place (X,Y,R)
- Laser line: Detect object position in a single work plane (X,Z)
- 2D+Laser: Detection of objects positioned in two or more layers (X,Y,Z,R)





3D:
- Stereo vision: Detect object position and orientation based binocular vision and image processing
- Infrared: Use one image to map a part in 3D
- Laser point cloud: Structured and sparse laser reflection measurements to identify objects or features in robot workspace

Force sensing:
Identify forces and torque in 3DOF

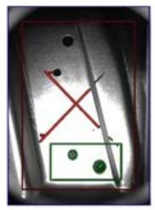
GM RESEARCH & DEVELOPMENT

9

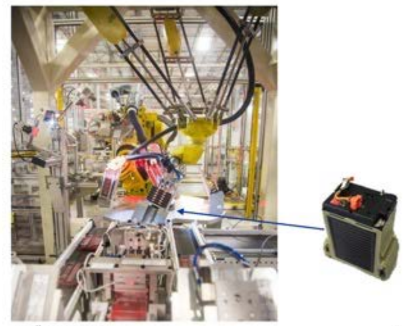
ASSEMBLY OPERATIONS

2D and 3D perception systems: Place parts in reference to a target feature (e.g.: holes) or surface matching

- **Body Assembly:** Sheet metal component with target features for positioning and alignments



- **Battery Assembly:** Metal and plastic components with locating features



media.gm.com

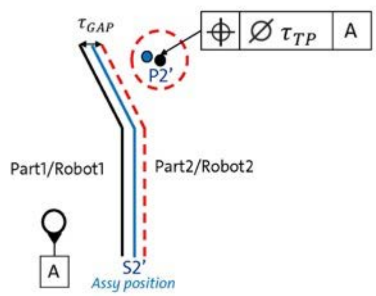
GM RESEARCH & DEVELOPMENT

| 8

VISION FOR PATH PLANNING

- **Offset from taught points**
Relationship between datum and key feature is taught and vision is used to define an offset

- **Direct pose programming**
Relationship between datum and key feature programmed based on the product requirements



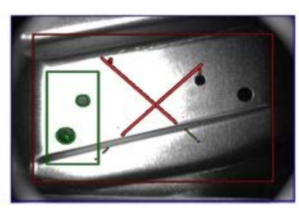
GM RESEARCH & DEVELOPMENT

| 10

CHALLENGES: IMAGE ACQUISITION AND PROCESSING

- **Part properties:** Poor contrast or reflection. Images are required to have good contrast between the work environment, background, target feature

- **Environmental conditions:** Changes in ambient lighting can produce inconsistent performance.



GM RESEARCH & DEVELOPMENT

| 12

CHALLENGES: IMAGE ACQUISITION AND PROCESSING

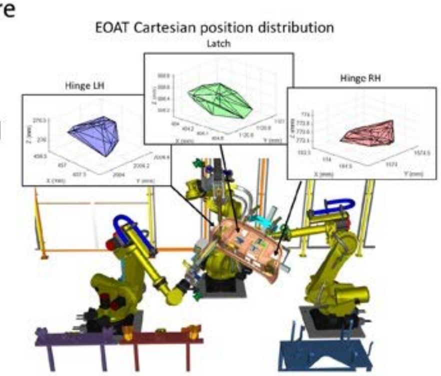
- **Field of View:** Narrow FOV requires multiple robot movements which increases operation cycle time
- **Accuracy:** Aggregated error of optical instruments, image processing and feature extraction algorithm and robot movements
- **Access:** Robot mounted sensors require might face access and constraints to get proper image

GM RESEARCH & DEVELOPMENT

| 11

CHALLENGES: INTEGRATION AND ROBOT ACCURACY

- **Integration:** Additional hardware and software requirements that increate system latency
- **Robot accuracy:** Inherent errors in robot mastering and calibration along with thermal expansion/contraction of robot arm affect accuracy of vision system and commanded positions



GM RESEARCH & DEVELOPMENT

| 13

OPPORTUNITIES: QUALITY

- **Accurate imaging:** Use solutions often used for metrology for accurate robot guidance
- **Accurate positioning:** Reduce the errors in robot positioning for both image acquisition and assembly
- **Post-process inspection:** Use vision system for robot guidance and inspection to enable “Built-in-quality” where no bad parts leave the cell

GM RESEARCH & DEVELOPMENT

| 15

OPPORTUNITIES: PRODUCTIVITY

- **Reduce robot movement:** Larger field of view would reduce the need to collect multiple images to command an absolute or relative position to one or multiple robots
- **Faster image acquisition:** Solutions robust to changes in part reflection, glare, or ambient light will prevent the system of having to take multiple images of a single feature

media.gm.com

GM RESEARCH & DEVELOPMENT

| 14

RESEARCH AND DEVELOPMENT REQUIREMENTS

- Extended Field of View: Increase the working distance to the observable world from a single or multiple camera locations
- Increase accuracy: Reduce or estimate error or vision system and robot arm
- Robust to different parts and work environment: Use of different wavelengths for perception
- Fast image acquisition: reduce time to acquire image and identify features

GM RESEARCH & DEVELOPMENT

| 16

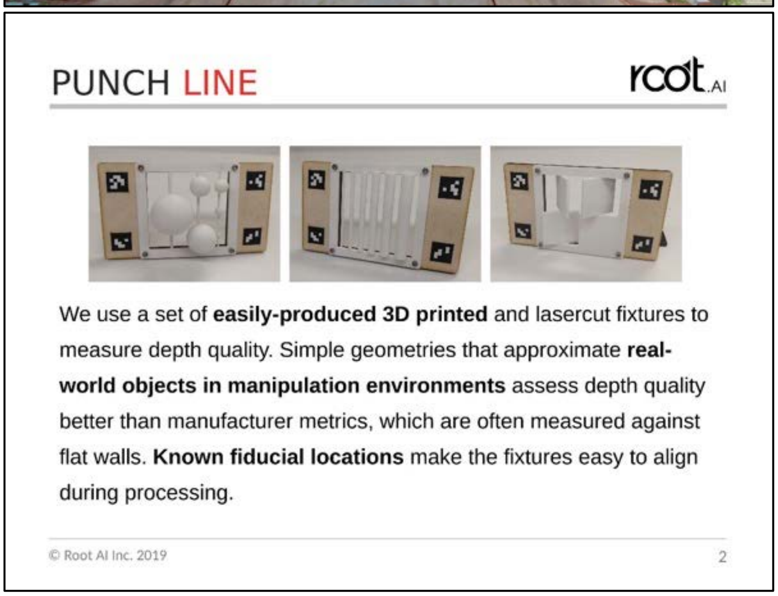
STANDARDIZATION REQUIREMENTS

- Define a common testing setup
- How to measure accuracy?
- How to differentiate between sensor, algorithm, and robot error?

GM RESEARCH & DEVELOPMENT

| 17

B.3. Depth Quality Assessment at Close Range Using 3D Printed Fixtures by Michele Pratusevich, Root AI



ROOT AI AND ME

© Root AI Inc. 2019 3

AGRICULTURAL ROBOTICS

© Root AI Inc. 2019 4

VIRGO THE HARVESTER

© Root AI Inc. 2019 5

TOMATOES: SMALL AND CLOSE

To harvest, we need accurate representations in 3D. USDA provides **weight** (but not size) regulations so our information was gathered from customers. Cherry / grape tomatoes are about **20 mm** in diameter. Rows are narrow (~1 meter wide) so any manipulator needs to reach between 30cm and 70cm to harvest tomatoes. Need to select a good enough camera.

 6

WHAT IS GOOD ENOUGH?

Intel Realsense publishes their testing methodology (RMSE of a white wall). Other manufacturers (ZED, Asus) don't.

https://www.intel.com/content/dam/support/us/en/documents/emerging-technologies/intel-realsense-technology/RealSense_DepthQualityTesting.pdf

RMSE on a white wall does not approximate our complex scene.

 7

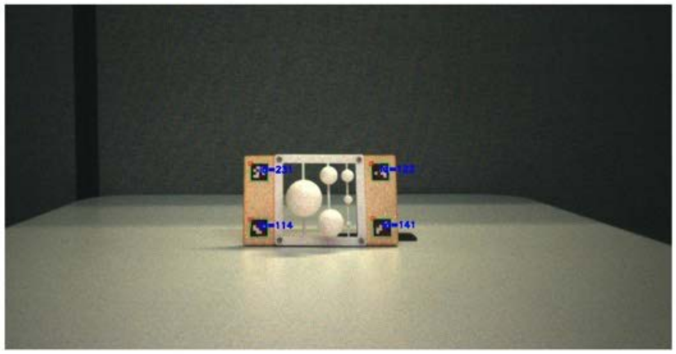
FIXTURE DESIGN

We use a set of **easily-produced 3D printed** and lasercut fixtures to measure depth quality. Simple geometries that approximate **real-world objects in manipulation environments** assess depth quality better than manufacturer metrics, which are often measured against flat walls. **Known fiducial locations** make the fixtures easy to align during processing.

 8

EVALUATION STEP 1

root.AI

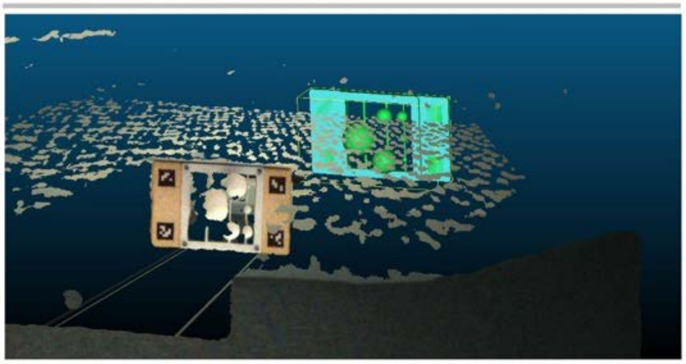


Capture RGB image and detect fiducials

9

EVALUATION STEP 2

root.AI

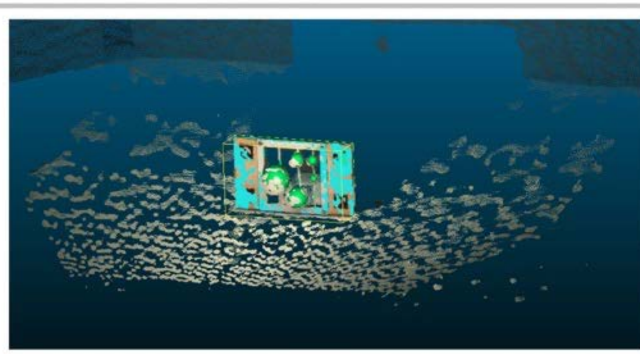


Capture pointcloud

10

EVALUATION STEP 3

root.AI



Align pointcloud to 3D model, crop, and evaluate

11

DEPTH METRICS

rcot.AI

RMSE of pixels in the test object segment (and not anywhere else).

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(o_i - e_i)^2}$$

Density of points that are on surfaces the camera should see.

$$\text{Density} = \frac{1}{A}\sum_{i=1}^{n}\begin{cases}1 & \text{if} \quad |o_i - e_i| < t \\ 0 & \text{otherwise}\end{cases}$$

$$A = \sum_{f \in \text{faces}}\begin{cases}\text{Area}(f) & \text{if} \quad \arccos(f_{\perp} \cdot c) > \frac{\pi}{2} \\ 0 & \text{otherwise}\end{cases}$$

 12

CAMERA SELECTION

rcot.AI

RMSE (m) AND DENSITY (PIXELS / m^2) FOR 3 CANDIDATE CAMERAS

Fixture	Metric	D415	D435	ZED Mini
Cylinders	RMSE	**0.00177**	0.00200	0.00319
	Density	0.00144	0.00137	**0.00197**
Spheres	RMSE	**0.00269**	0.00415	0.00532
	Density	0.00150	0.00098	**0.00182**
Angled plates	RMSE	**0.00223**	0.00286	0.00324
	Density	0.00145	0.00140	**0.00223**

For our application, the **D415** gives the best accuracy.

 13

SANITY CHECK

rcot.AI

When the target is farther from the camera, the RSME is higher and the density is lower. The target was aligned as close to the center of the FOV as possible.

RMSE (m) AND DENSITY (PIXELS / m^2) FOR D415 WITH CYLINDERS AT VARIOUS DISTANCES (INCHES)

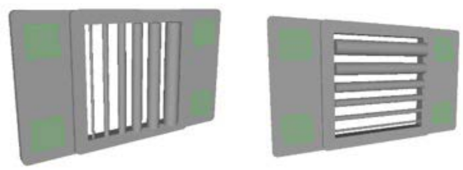
Distance	RMSE	Density
16	1.64	3.36
20	1.97	1.61
24	1.76	1.44
28	2.18	0.91
32	1.97	0.79
36	1.89	0.61
40	2.05	0.48

 14

USES AND EXTENSIONS

3D printing targets makes creating standard targets easy to reproduce. More work is needed to determine optimal materials, settings, etc. Can create targets specific to different characteristics:

- Shapes
- Sizes
- Surface colors / textures
- Specularity



 15

LEARN MORE

- Github: http://github.com/root-ai/depth-quality
- More about Root AI: http://root-ai.com
- Email me: mprat@root-ai.com
- Paper on ArXiv:

 16

B.4. Depth Camera Image Quality Definition and Measurement by John Sweetser, Intel

DEPTH QUALITY DEFINITION AND MEASUREMENT

John Sweetser

RealSense Group, Intel Corp

DISCUSSION TOPICS

- WHAT IS DEPTH QUALITY AND WHAT MAKES A GOOD DEPTH CAMERA?
  - QUANTITATIVE & QUALITATIVE ASSESSMENT
- DEPTH QUALITY EVALUATION
  - PERFORMANCE METRICS (KPI'S)
  - TEST AND CHARACTERIZATION METHODS AND TOOLS
  - PERFORMANCE CRITERIA (STANDARDS)
- SAMPLE DEPTH QUALITY TEST DATA
- FACTORS THAT INFLUENCE DEPTH QUALITY
- KPI'S NOT CURRENTLY TESTED
- EFFECT OF CAMERA CALIBRATION ON DEPTH QUALITY

2

WHAT IS DEPTH IMAGE QUALITY?

- **Depth Camera:** adds the distance (Z) dimension to traditional 2D RGB or B&W image. *attributes such as sharpness, distortion/uniformity, color fidelity, noise, and dynamic range, etc.*
- **Depth Image:** typically represented as 2D 'depth map' or 3D 'point cloud'.
- **Depth Image Quality:** Evaluation of the depth image quantitatively (using predefined metrics) or qualitatively (using visual clues).

Typically, quantitative metrics are used in simplified scenes and controlled conditions and qualitative assessment is used in arbitrary or complex scenes

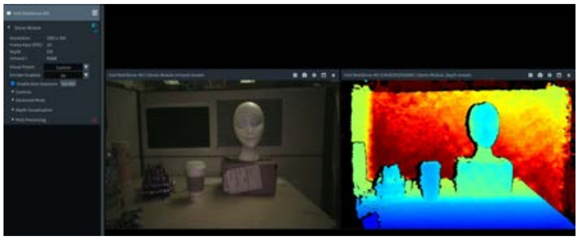
intel | 3

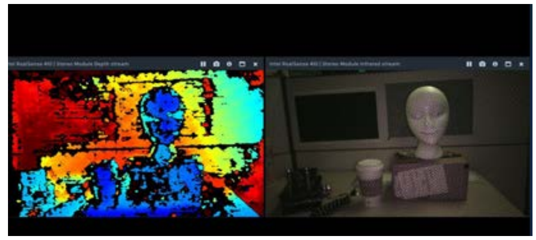
WHAT MAKES AN IDEAL DEPTH CAMERA?

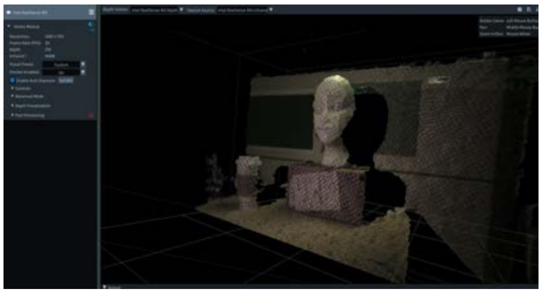
1. **See everything:**
   a) All conditions: From darkness to bright sunlight
   b) All materials & objects
   c) All ranges
   d) No interference
2. **See it with little noise (high precision)**
3. **Get exact distance (high accuracy)**
4. **And cheap, small, low-power, wide field-of-view, high-speed, color…**

intel | 4

SAMPLE "GOOD" AND "BAD" DEPTH IMAGES







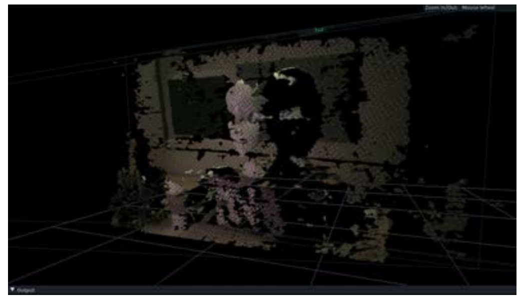
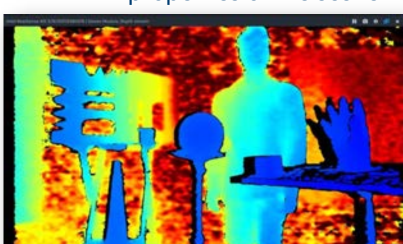
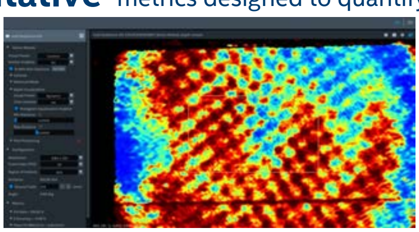


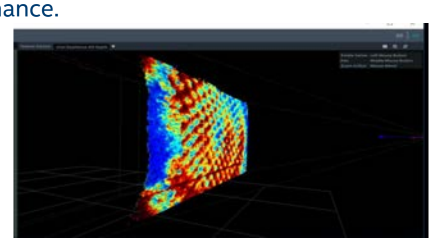
IMAGE QUALITY EVALUATION

Qualitative Based on visual inspection of the depth map or point cloud and assess quality based on known properties of the scene



Quantitative Based on measurements performed on the camera, depth data can be analyzed to produce metrics designed to quantify performance.





KEY DEPTH METRICS

A very good snapshot of depth performance can be seen from the histogram of flat target depth values. Ideally, this distribution is narrow, centered near the known (ground truth) distance, and has a complete number of sample points.

- **Fill Ratio:** Percentage of "valid" (w/non-zero depth) pixels over ROI.

*Typical good value: >99%, <98% poor

- **Z-Accuracy:** Offset of mean/median depth from ground truth.

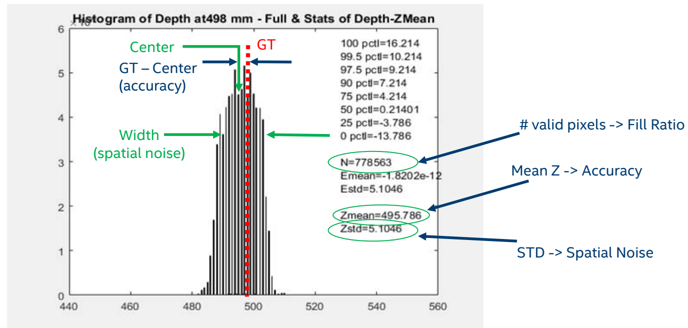
Typical good value: <1%, >2% poor

- **RMS Error (Spatial Noise):** Variation in depth over ROI.

Typical good values: <0.4% (~0.2 pixels), >1% poor

- **Temporal Noise:** Variation in depth per pixel over time (frame-to-frame).

Typical good values <0.2%, >0.5% poor



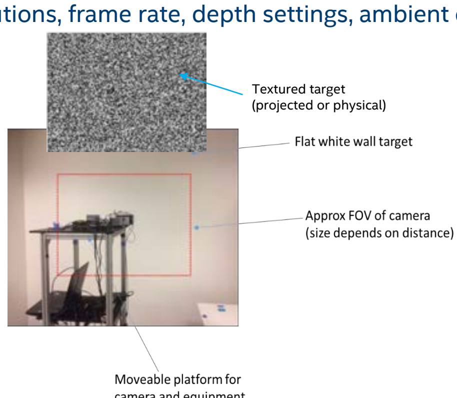
Sample of raw depth data from a DS5 camera during a flat target test

*For D415 @ Z ~ 1m, HD resolution, center 40% ROI, Active
Consult datasheet for latest specifications

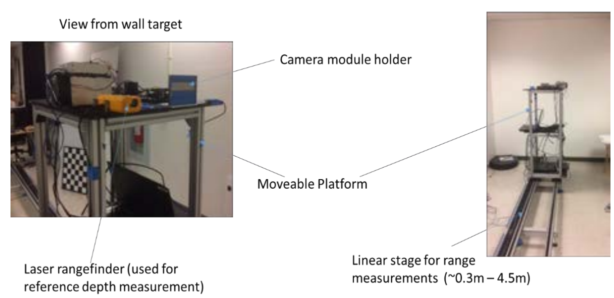
Metric values may be expressed in absolute units, e.g., mm or as % of depth.

IMAGE QUALITY MEASUREMENT – BASIC Z-PERFORMANCE

- Quantitative depth quality is evaluated primarily based on "flat target" testing (either textureless or textured). This provides a simple, well-defined, and standardized environment to capture images and compute metrics.
- In all measurement methods, image data is captured and then analyzed, either off-line or in real-time, to compute the performance metrics.
- Measurements are performed as a function of distance from the target and may be run for different resolutions, frame rate, depth settings, ambient conditions, etc.

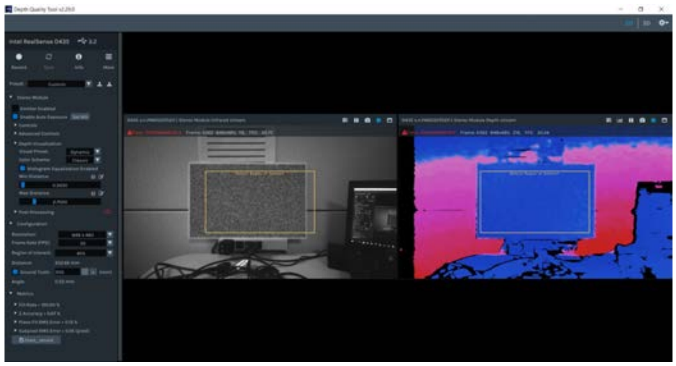


A typical camera characterization apparatus



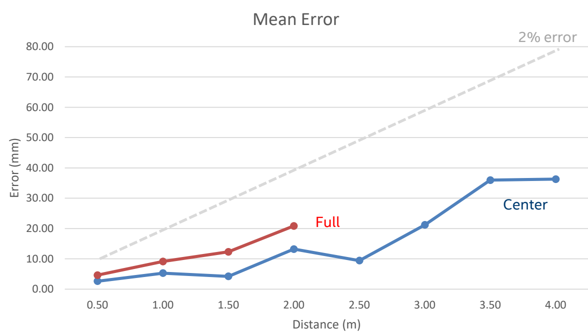
DEPTH QUALITY MEASUREMENT TOOLS

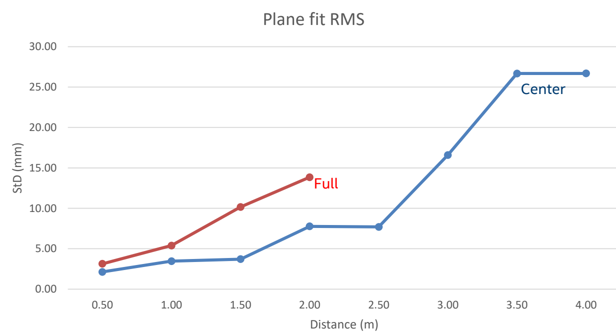
- **Offline**: Typically used for official validation. Depth data captured and analyzed later.
- **Real-time**: Test application that captures, analyzes data, and computes metrics in real-time (per frame or based on sequence). Metrics are usually a subset of full validation that contain only the key metrics needed for basic depth camera health check.
- For D400 cameras, **Depth Quality Tool** is the recommended tool for basic real-time measurements



SAMPLE DEPTH QUALITY DATA – ASR SHORT RANGE

1280x720, 30 FPS, P=210mW, AE
Target: Flat white wall, ~100-200 Lux fluorescent lighting





mean Error (mm)								
SN	0.50	1.00	1.50	2.00	2.50	3.00	3.50	4.00
Center 40%	2.62	5.28	4.21	13.19	9.41	21.20	35.95	36.28
Full FOV	4.60	9.12	12.29	20.85				

std (mm)								
SN	0.50	1.00	1.50	2.00	2.50	3.00	3.50	4.00
Center 40%	2.13	3.46	3.71	7.77	7.69	16.59	26.68	26.69
Full FOV	3.13	5.39	10.16	13.85				

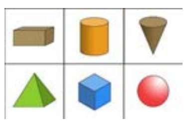
FACTORS AFFECTING DEPTH QUALITY

Product use cases largely drive environment in which depth quality is assessed from among factors below

Lighting - should be tested in different lighting conditions in which the product is used. Different technologies will behave differently in specific lighting conditions.

Different Materials, Flat uniform surfaces (Ex. White board/wall), Textured Patterns
- Ex: Autonomous vacuum cleaners would test different floor materials such as light and dark tile, wood, carpet, and linoleum; body scanning would test materials that might be worn by the user and different colors/patterns of that material.

Range or Distance
- Ex 1: A room scanning device designed to remain in the center of a large space and rotate while capturing walls, ceiling, and objects in detail will require accurate depth at long distance.
- Ex 2: A robot or drone in motion can use depth at longer distances for path planning without requiring accurate depth at those distances.



Shape
- Quantitative testing currently done with flat targets due to Ground Truth availability.
- Qualitative testing looking at the point cloud for edge fidelity, flat or round surfaces, and proper angles on different geometric shapes.

ADDITIONAL 3D IMAGE QUALITY CHARACTERISTICS OF INTEREST

Types of quantitative testing/characterization not currently done (@ Intel)

- **2D (x,y) spatial resolution:**
Resolution chart with variable width slots or features.

- **Minimum detectable object size:**
Targets with variable size objects (spheres, cylinders) - 3D resolution.

- **Edge Fidelity:**
Sharpness of edges (depth discontinuities).

- **Full 3D Object/Scene Reproduction:**
Error in reproducing a specific scene or object (e.g., mannequin). RMSE from ground truth.

EFFECT OF CALIBRATION ON DEPTH QUALITY

- **Approximately 27 parameters that are determined during a full calibration procedure:**
  - **Intrinsic – individual camera factors (PP, FL, distortion)**
  - **Extrinsic – relative left-right camera position & orientation**

- **Post-factory re-calibration can usually be done by adjusting 1-2 intrinsic and/or extrinsic parameters:**
  - **Px/Py – shift of lens-sensor position to adjust disparity or alignment of images (~0.2 – 1 pixel).**
  - **Rx/Ry/Rz – rotation of camera for same purpose (<0.2 deg).**

Most degradation of depth quality can be corrected quickly with proper adjustment of one or more of these parameters.

Temperature variations, mechanical shock/vibration or stress can lead to degradation that requires re-calibration.

intel | 13

intel

BACKUP

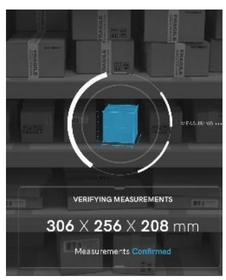
OTHER PERFORMANCE METRICS

"Fill Factor" – Combines fill ratio, Z accuracy, and RMSE in to a single figure of merit.
Example: % of all pixels that are within 3% of ground truth distance.

3D CAMERA APPLICATIONS

Robotics

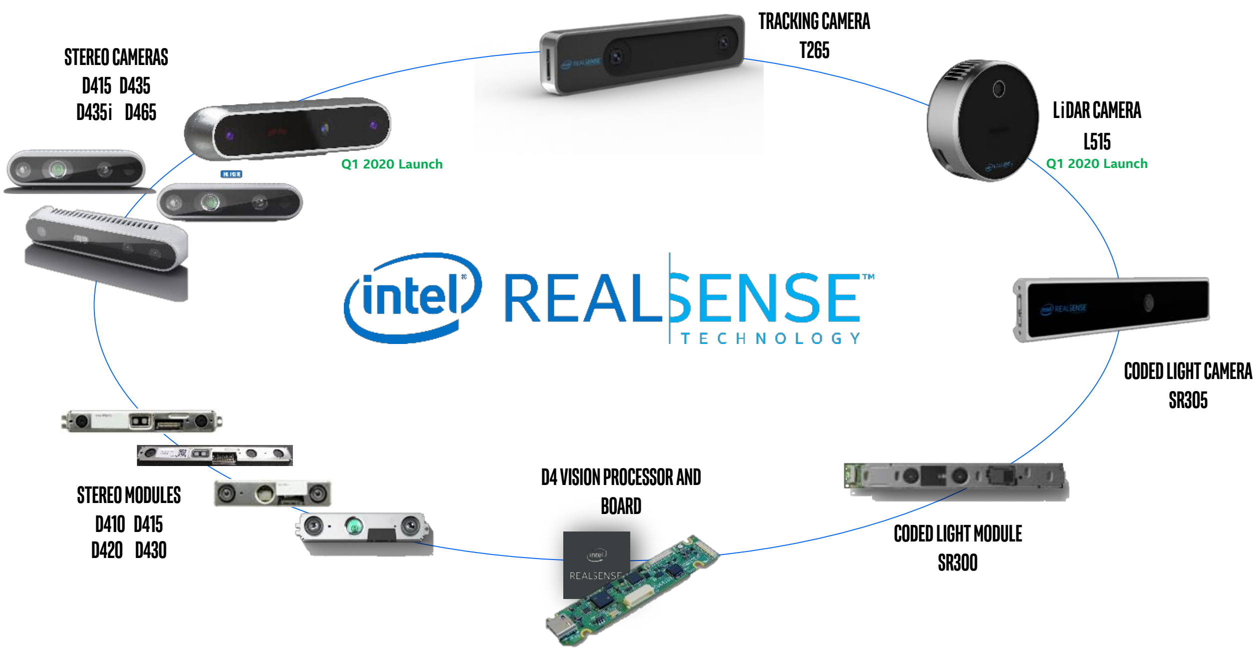
Scanning



Measurement

Tracking

Facial Authentication

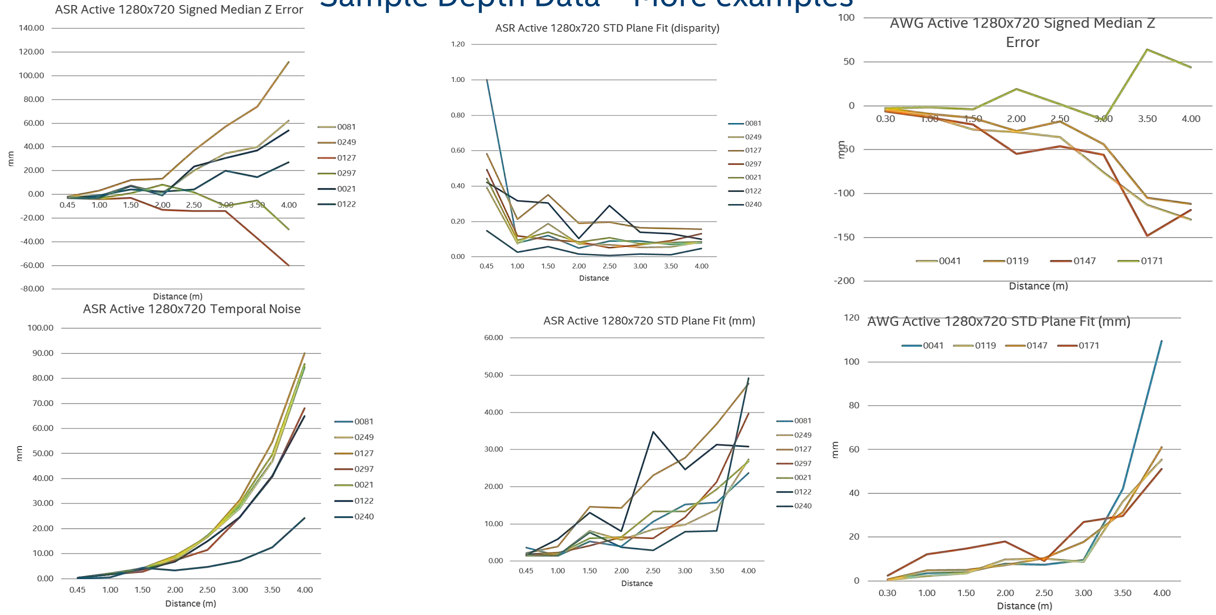


DEPTH QUALITY DEPENDENCE ON TECHNOLOGY

- Depth quality evaluation methods and metrics in general do **NOT** depend on underlying technology, however the image quality itself may have technology-specific characteristic.
- Active depth Technologies such as Structured Light and TOF rely on projected light and therefore work well in scenes with ***little texture*** and ***low lighting***, such as uniform walls in a factory or office environment. Therefore, these are the conditions recommended for evaluation.
- Stereo depth (such as D400 family) which does not rely solely on projected light and can benefit from natural texture and ambient lighting, may be evaluated in a variety of scenes and conditions.

	Lighting			Target Scene				Distance	
	Low Light	Sunlight	Indoor normal light	Uniform, high reflectivity surface (Ex. Flat White wall)	Texture	Geometric shapes/edges	Materials (Low Reflectivity, Diffuse, Dark)	Near	Far
Stereo									
Active Stereo									
TOF									
Structured Light									

Sample Depth Data – More examples



INTEL® REALSENSE™ D400 SERIES DEPTH QUALITY

- RealSense™ 400 series provides excellent depth quality under all lighting conditions, and longer range
- Great configurability - Viewer and Depth Quality tools in SDK provide different Presets (High Density, High Accuracy, Close Range, Hand etc.) OR users can tune their own for their applications

2D RGB Image

HIGH DENSITY PRESET

D415

DEPTH MAP(1280x720)

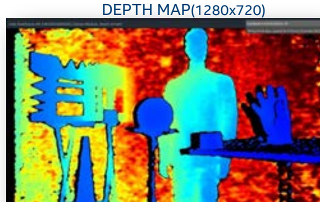
POINTCLOUD WITH COLOR TEXTURE

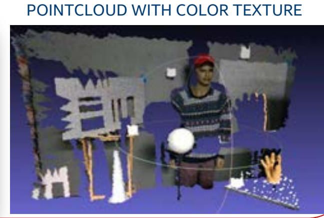
Ex. Usages:

BGS

3D Enhanced Photography

Collision Avoidance

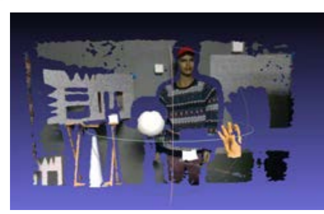




HIGH ACCURACY PRESET

Ex. Usages:

3D Object Scan

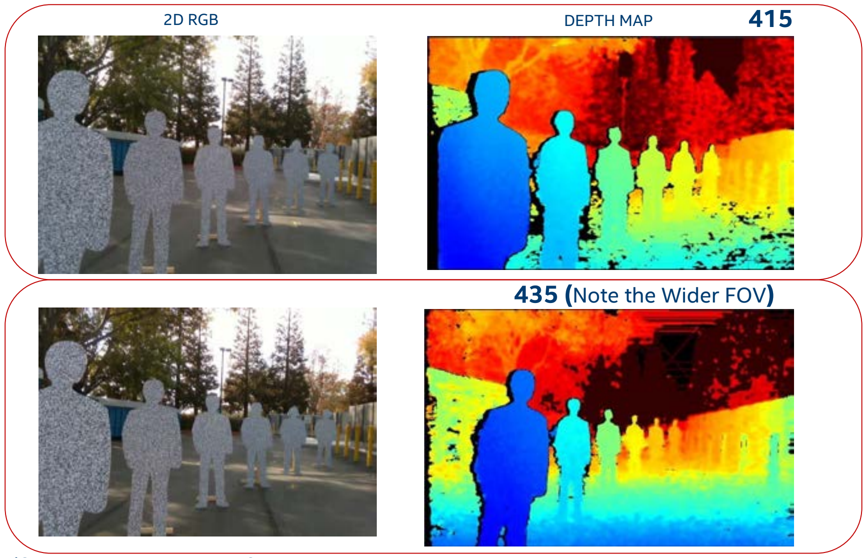


3D test scene configured to include a variety of object types, textures and distances, captured through RealSense™ Viewer 2.8.1.

Intel RealSense Group

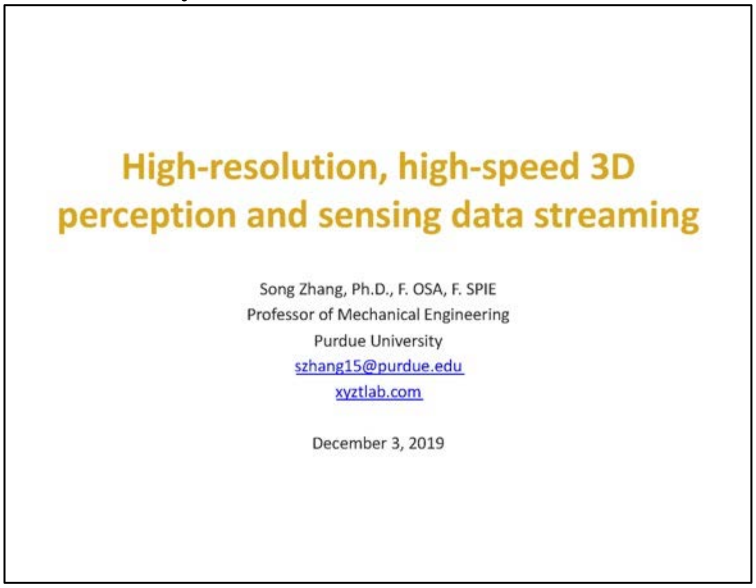
INTEL® REALSENSE™ D415/D435 OUTDOORS

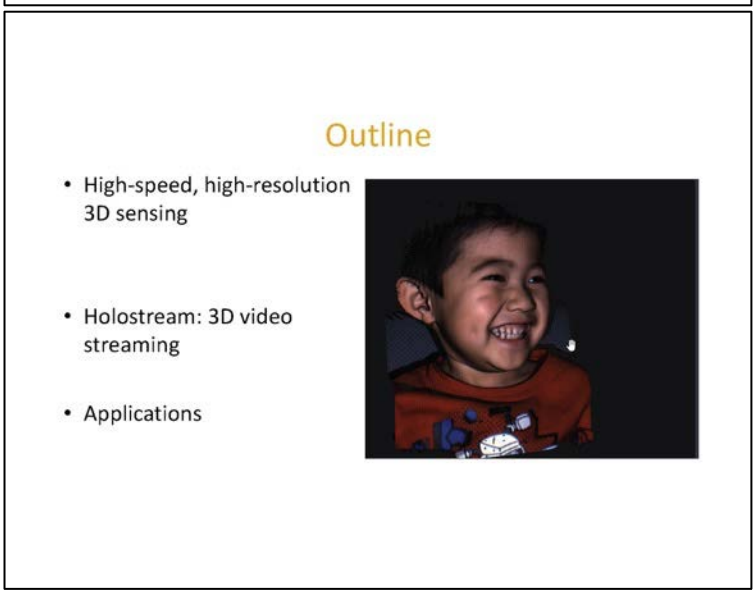
- Stereo takes advantage of visible light for best Outdoor performance and Range
- Projector can be off with enough visible light and texture => low power!



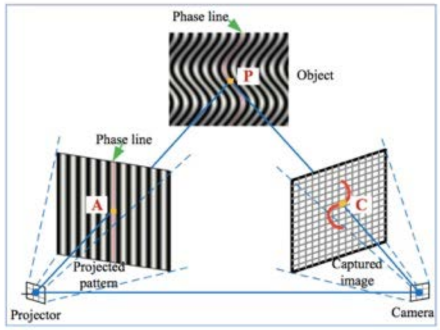
3D test scene outdoors captured through RealSense™ Viewer 2.8.1 – 10m?.

B.5. High-resolution, high-speed 3D perception and sensing data streaming by Song Zhang, Purdue University





Structured light technology



Three-step phase shifting algorithm

- Phase shifted fringe images

$$I_1(x,y) = I'(x,y) + I''(x,y)\cos[\phi(x,y) - 2\pi/3]$$
$$I_2(x,y) = I'(x,y) + I''(x,y)\cos[\phi(x,y)]$$
$$I_3(x,y) = I'(x,y) + I''(x,y)\cos[\phi(x,y) + 2\pi/3]$$

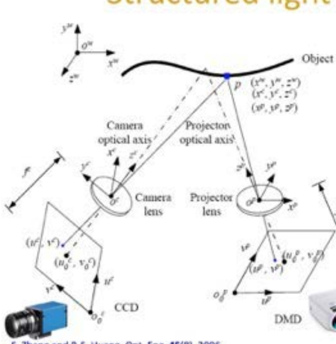
- Wrapped phase

$$\phi(x,y) = \tan^{-1}\frac{\sqrt{3}[I_1(x,y) - I_3(x,y)]}{2I_2(x,y) - I_1(x,y) - I_3(x,y)}$$

- 2D texture

$$I'(x,y) = [I_1(x,y) + I_2(x,y) + I_3(x,y)]/3$$

Structured light system calibration



S. Zhang and P. S. Huang, Opt. Eng. **45**(8), 2006.

- Perspective transformation

$$s\{u,v,1\}^T = A[R,t]\{x^w, y^w, z^w, 1\}^T$$

- Extrinsic parameters

$$R = \begin{bmatrix} r_{00} & r_{01} & r_{02} \\ r_{10} & r_{11} & r_{12} \\ r_{20} & r_{12} & r_{22} \end{bmatrix}, \quad t = \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix}$$

- Intrinsic parameters

$$A = \begin{bmatrix} \alpha & \gamma & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

α, β: **focal lengths**

γ: **skew effect**

(u_0, v_0): **principal point**

3D coordinate calculation

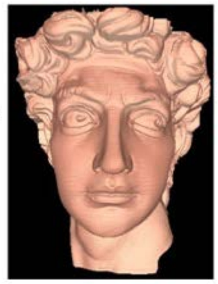
- From world to camera image coordinates

$$s^c \begin{Bmatrix} u^c \\ v^c \\ 1 \end{Bmatrix} = \begin{bmatrix} \alpha^c & \gamma^c & u_0^c \\ 0 & \beta^c & v_0^c \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{00}^c & r_{01}^c & r_{02}^c & t_x^c \\ r_{10}^c & r_{11}^c & r_{12}^c & t_y^c \\ r_{20}^c & r_{21}^c & r_{22}^c & t_z^c \end{bmatrix} \begin{Bmatrix} x^w \\ y^w \\ z^w \\ 1 \end{Bmatrix}$$

- From world to projector image coordinates

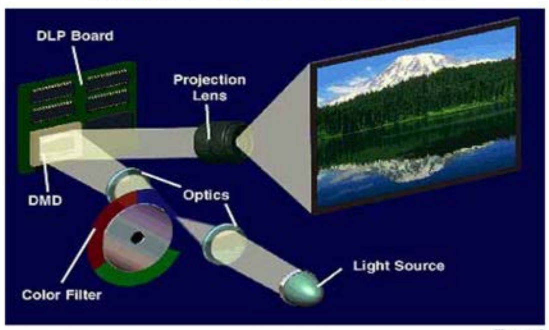
$$s^p \begin{Bmatrix} u^p \\ v^p \\ 1 \end{Bmatrix} = \begin{bmatrix} \alpha^p & \gamma^p & u_0^p \\ 0 & \beta^p & v_0^p \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{00}^p & r_{01}^p & r_{02}^p & t_x^p \\ r_{10}^p & r_{11}^p & r_{12}^p & t_y^p \\ r_{20}^p & r_{21}^p & r_{22}^p & t_z^p \end{bmatrix} \begin{Bmatrix} x^w \\ y^w \\ z^w \\ 1 \end{Bmatrix}$$

- Absolute phase constraint: $u^p = f[\Phi_a(u^c, v^c)]$
- 7 equations
- 7 unknowns: (x^w, y^w, z^w), u^p, v^p, s^c, s



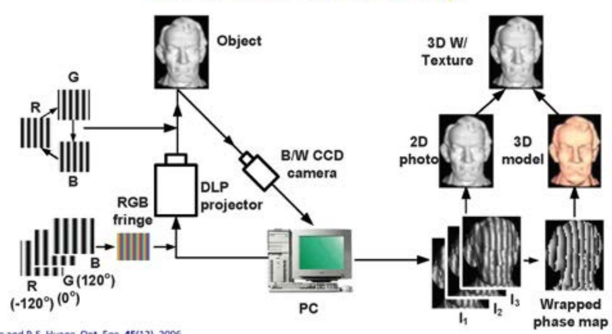
S. Zhang and P. S. Huang, Opt. Eng. **45**(8), 2006.

Single-chip DLP projector



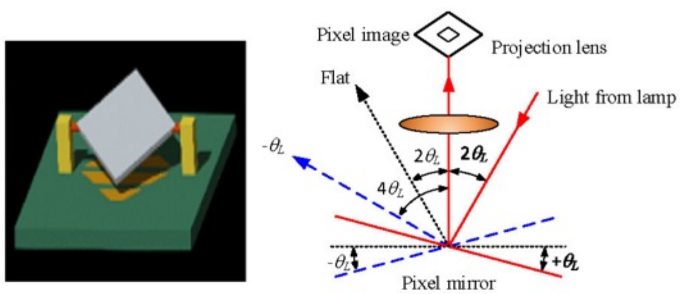
Pictures from www.ti.com

Real-time 3D sensing



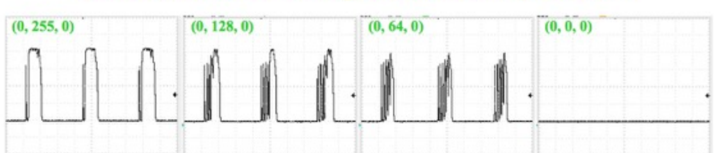
S. Zhang and P. S. Huang, Opt. Eng. **45**(12), 2006

DLP technology



Pictures from www.ti.com

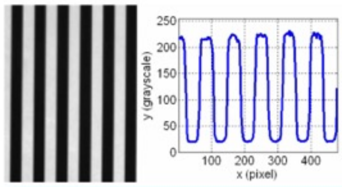
Limitations of using sinusoidal patterns

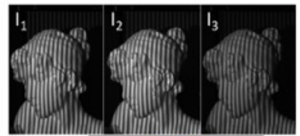


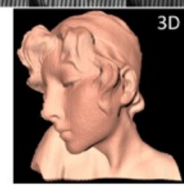
Projected timing signals with different grayscale input

- Precise synchronization requirement
- Speed limit of 120 Hz
- Projector's nonlinear gamma effect

Binary defocusing method





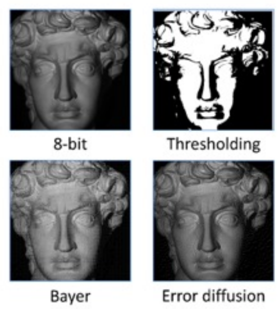


- DLP Discovery 4100 (0.7")
  - Resolution: 1024 X 768
  - 8-bit image switching rate: 291 fps
  - 1-bit binary image switching rate: **32,552** fps

S. Zhang, Opt. Lett. **35**(7), 2010; S. Lei and S. Zhang, Opt. Lett. **34**(20), 2009

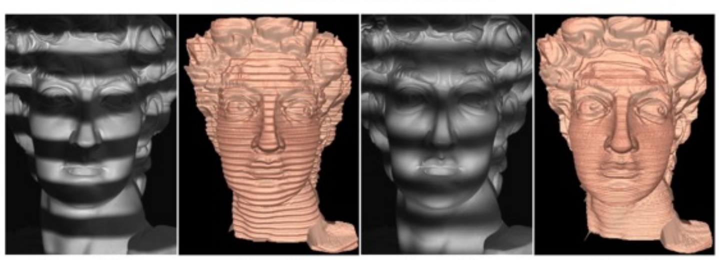
Dithering/halftoning

- Dithering (halftoning)
  - Approximate an image with fewer colors or bits
  - Adopted extensively in printing (halftoning)
- Methods
  - Single thresholding
  - Random dithering
  - Ordered dithering (Bayer, 1973)
  - Error-diffusion dithering (Floyd & Steinberg, 1976; Stucki, 1981)



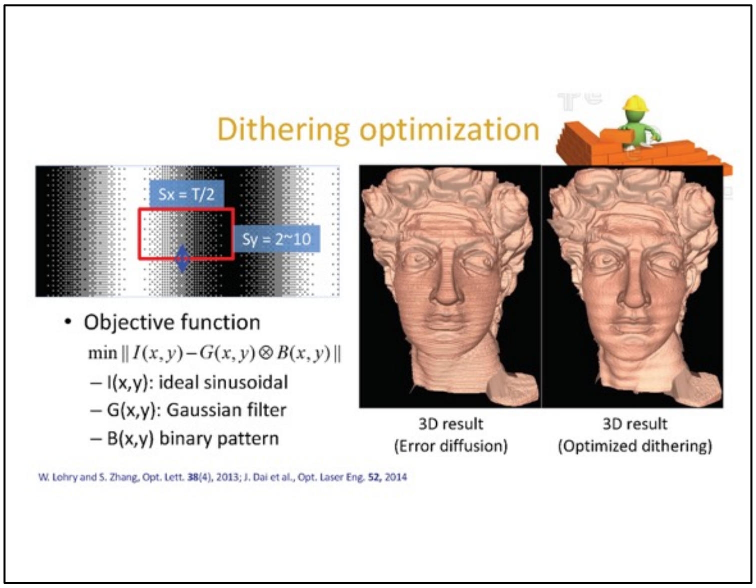
Y. Wang and S. Zhang, Appl. Opt. **51**(27), 2012

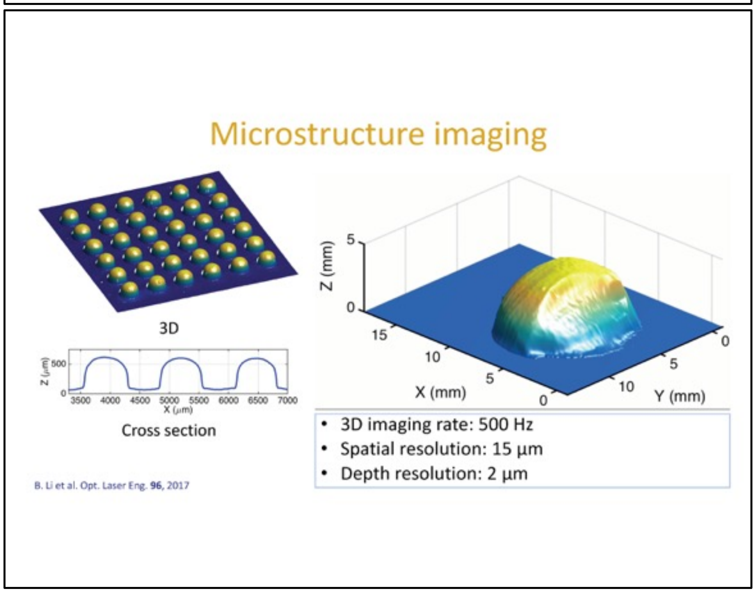
Comparing results

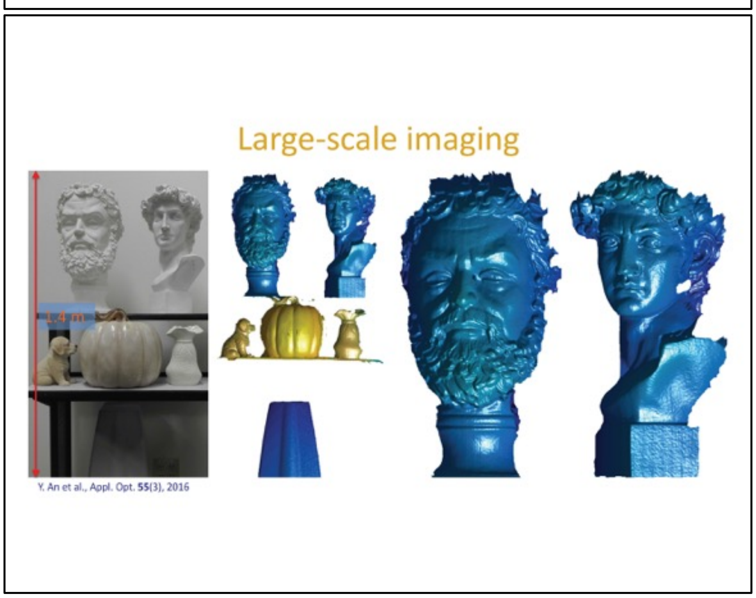


Fringe pattern (Square binary)	3D result (Square binary)	Fringe pattern (Error diffusion)	3D result (Error diffusion)

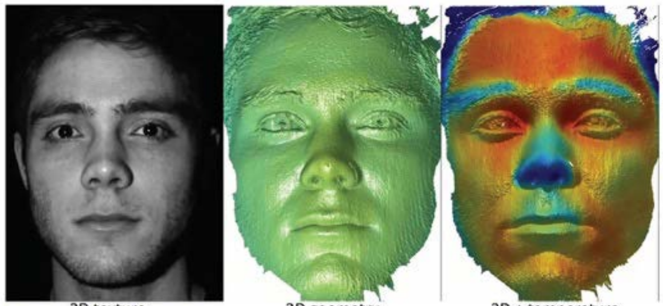
B. Li, et al., Opt. Laser Eng. **54**, 2014


Dithering optimization
Sx = T/2
Sy = 2~10
• Objective function
$\min \| I(x,y) - G(x,y) \otimes B(x,y) \|$
– I(x,y): ideal sinusoidal
– G(x,y): Gaussian filter
– B(x,y) binary pattern
3D result
(Error diffusion)
3D result
(Optimized dithering)
W. Lohry and S. Zhang, Opt. Lett. 38(4), 2013; J. Dai et al., Opt. Laser Eng. 52, 2014


Microstructure imaging
3D
Cross section
Z (mm)
X (mm)
Y (mm)
• 3D imaging rate: 500 Hz
• Spatial resolution: 15 μm
• Depth resolution: 2 μm
B. Li et al. Opt. Laser Eng. 96, 2017


Large-scale imaging
1.4 m
Y. An et al., Appl. Opt. 55(3), 2016

Multimodal imaging



2D texture 3D geometry 3D + temperature

Y. An and S. Zhang, Opt. Express **24**(13), 2016

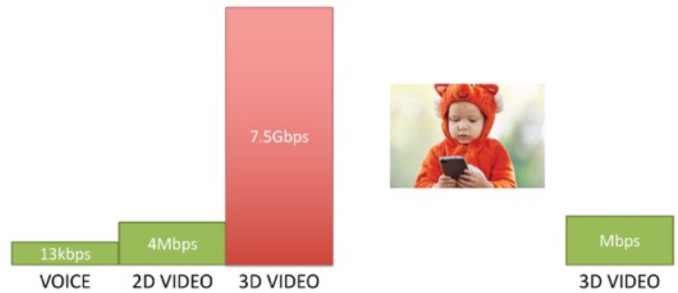
Holostream: 3D video communication

Smart phones with 3D cameras

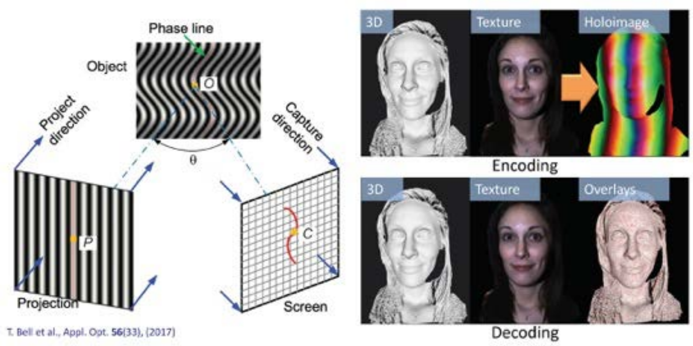
Apple Samsung Huawei Xiaomi

Over 100M phones with a 3D sensor shipped in 2018

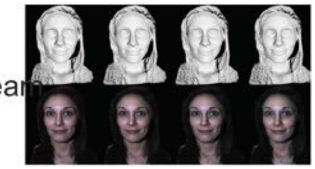
Why not 3D video communication?

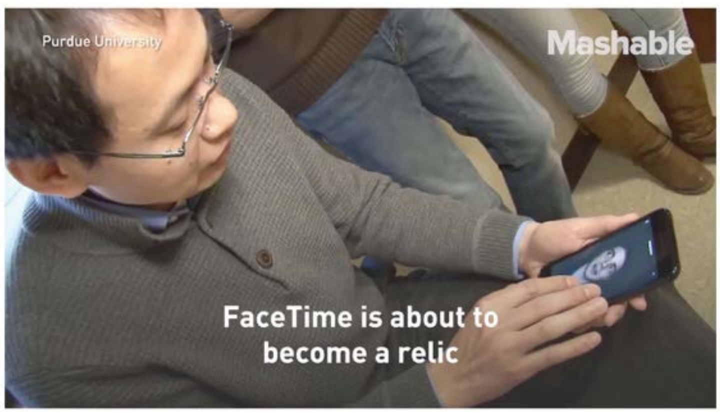


3D compression method



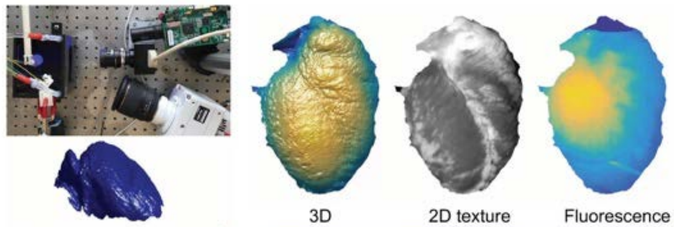
Holostream

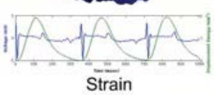




Applications

Cardiac imaging

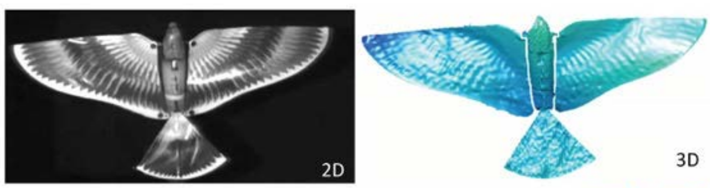




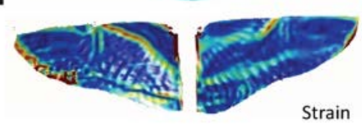
- Rabbit heart rate: 180 beats/min
- 3D imaging rate: 2,000 Hz
- Image resolution: 800 x 600

J. Laughner et al.,Heart and Cir Physio **303**(5), 2012; Y. Wang et al. Opt. Express, 21(5), 2013

Flapping wing robot

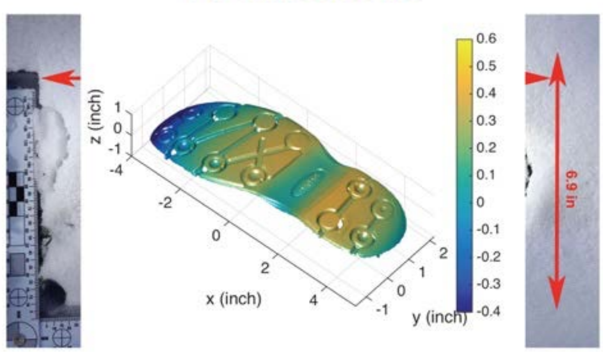


- Flapping rate: 21 cycles/sec
- 3D imaging rate: 5,000 Hz
- Resolution: 800 x 600

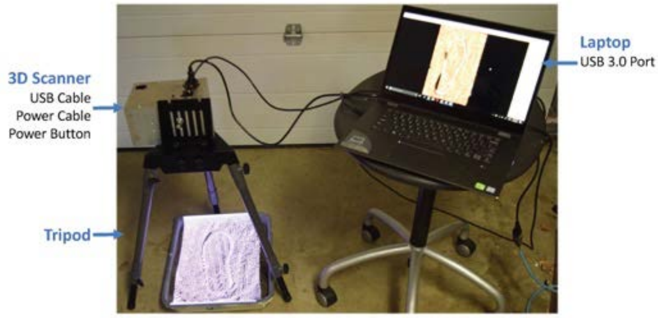


B. Li and S. Zhang, Meas. Sci. Technol. **29**(4), 2018

Forensic science

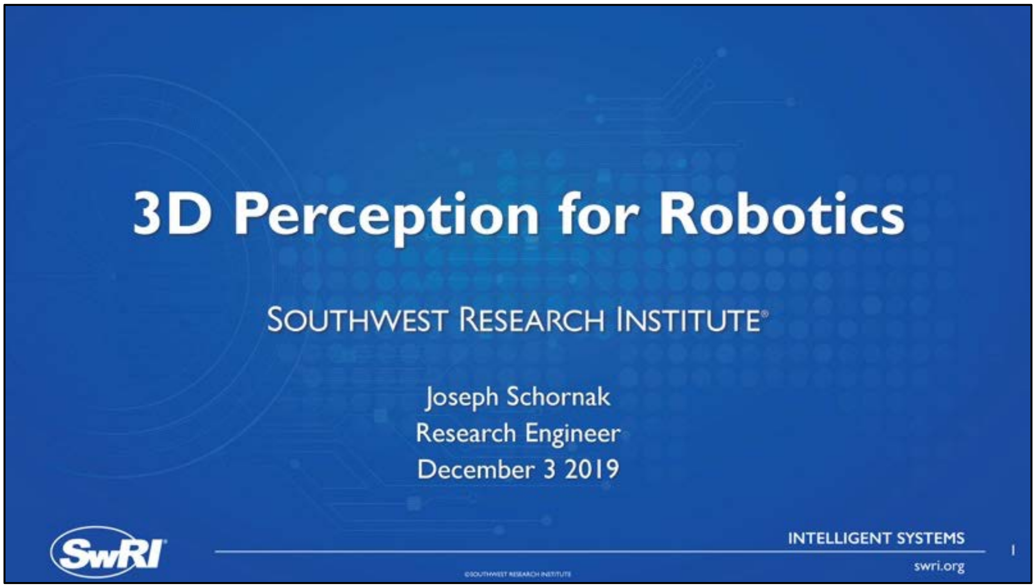


Forensic science



Forensic science

B.6. 3D Calibration and Perception for Robotic Scan-and-Plan Applications by Joseph Schornak, Southwest Research Institute



Intelligent Systems Division

- Autonomous vehicles
- High-reliability systems
- Traffic management systems
- Industrial automation

ActiveITS

SwRI INTELLIGENT SYSTEMS swri.org 3

Manufacturing and Robotic Technologies Department

- Advanced perception and planning for robotic applications.
- Industrial automation and controls.
- Systems incorporating both custom and off-the-shelf hardware.

SwRI INTELLIGENT SYSTEMS swri.org 4

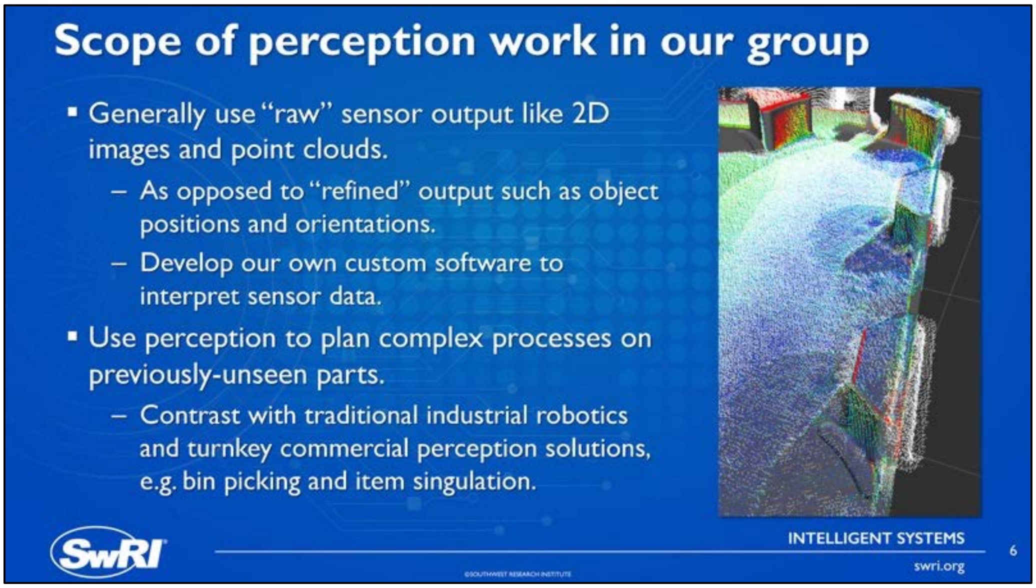
ROS-Industrial

- Goal is to develop software within the Robot Operating System (ROS) ecosystem targeted towards industrial applications.
- Consortium of companies and research groups provides funding.
- Resulting projects released as open-source repositories.

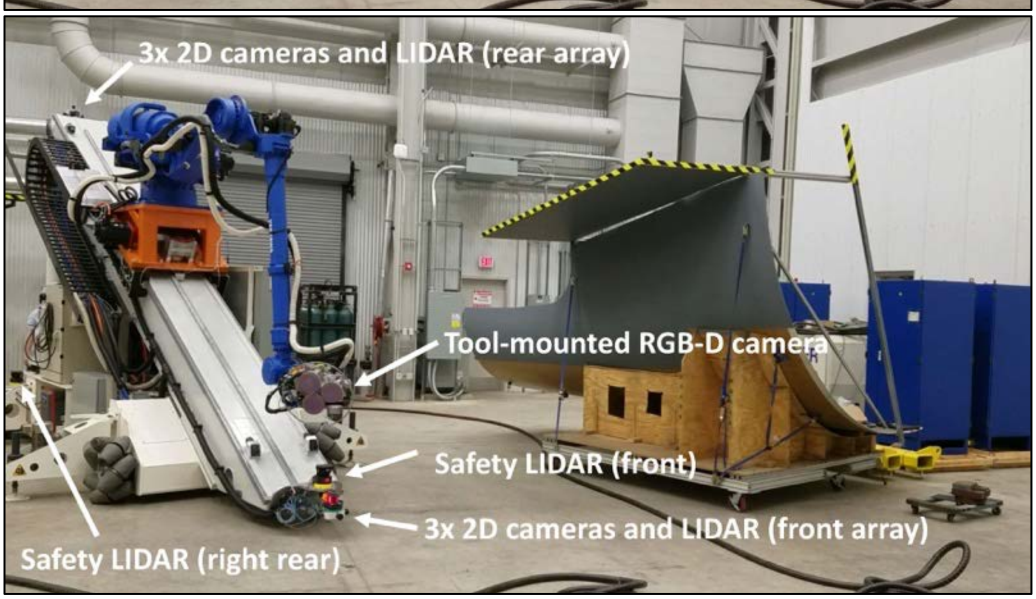
rosindustrial.org
github.com/ros-industrial
github.com/ros-industrial-consortium

ROS® industrial

SwRI INTELLIGENT SYSTEMS swri.org 5


Scope of perception work in our group
▪ Generally use "raw" sensor output like 2D images and point clouds.
– As opposed to "refined" output such as object positions and orientations.
– Develop our own custom software to interpret sensor data.
▪ Use perception to plan complex processes on previously-unseen parts.
– Contrast with traditional industrial robotics and turnkey commercial perception solutions, e.g. bin picking and item singulation.
SwRI
INTELLIGENT SYSTEMS
swri.org
6


A5: Advanced Automation for Agile Aerospace Applications
7

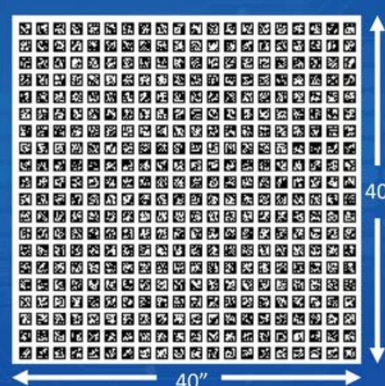

3x 2D cameras and LIDAR (rear array)
Tool-mounted RGB-D camera
Safety LIDAR (front)
3x 2D cameras and LIDAR (front array)
Safety LIDAR (right rear)

Development of open-source calibration tools

- Intrinsic calibration
  - Calculate lens optical parameters and distortion coefficients
- Extrinsic calibration
  - Solve 3D transforms to relate sensor to the robot and its surroundings.
  - Camera-to-tool, camera-to-world, robot kinematics, etc.
- Industrial Calibration
  - https://github.com/ros-industrial/industrial_calibration
- robot_cal_tools
  - https://github.com/Jmeyer1292/robot_cal_tools

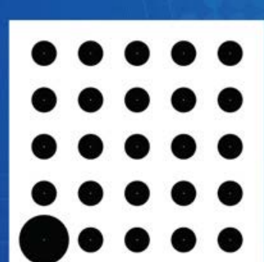
ArUco gridboard for calibrating in-hand sensors

- 20x20 array of squares gives 1600 corner features per board.
- Unique marker IDs allow use of partial target views for camera calibration.
- Big target fills camera field of view at practical working distance.



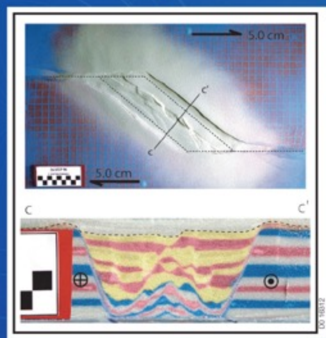
Modified OpenCV circle-grid target works too

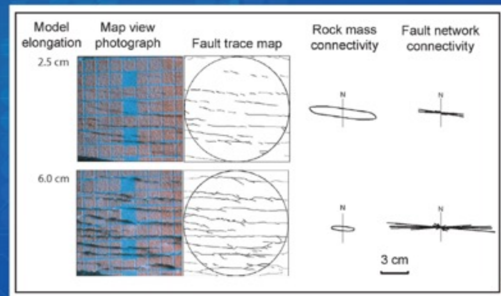
- Circle centroids are more accurate than square corners.
- Large corner dot denotes origin.



3D scans for physical analog geology simulation

- Enhance current analysis methods with dense, spatially-accurate structured light scans.

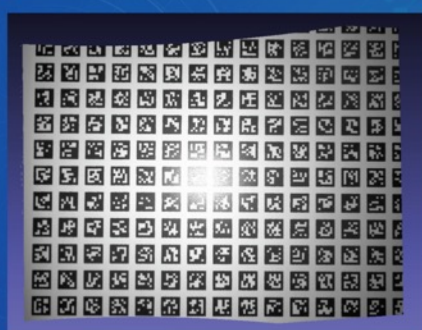


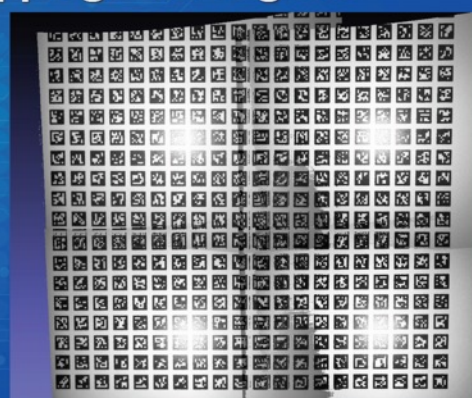


Analysis of two deformation stages from extensional fault system development illustrating evolution of rock mass connectivity and fault network connectivity. These are essential elements in quantifying permeability anisotropy in fractured and faulted reservoirs.

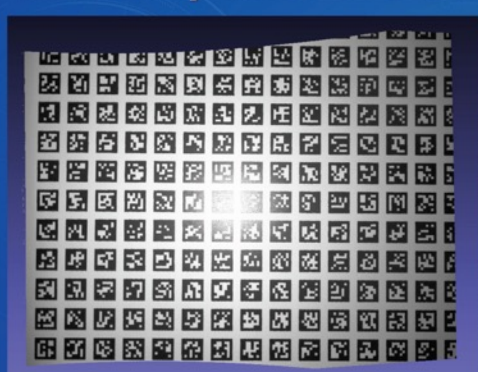
Figures from https://www.swri.org/physical-analog-modeling

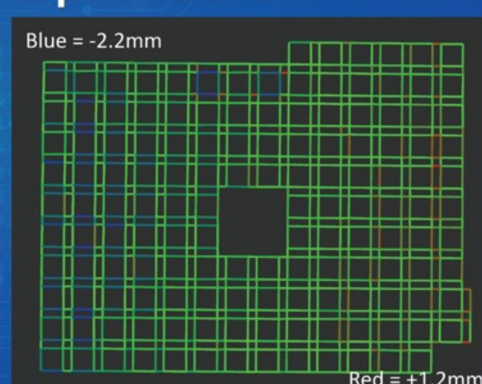
System extrinsic calibration scans show misalignment in overlapping scan regions

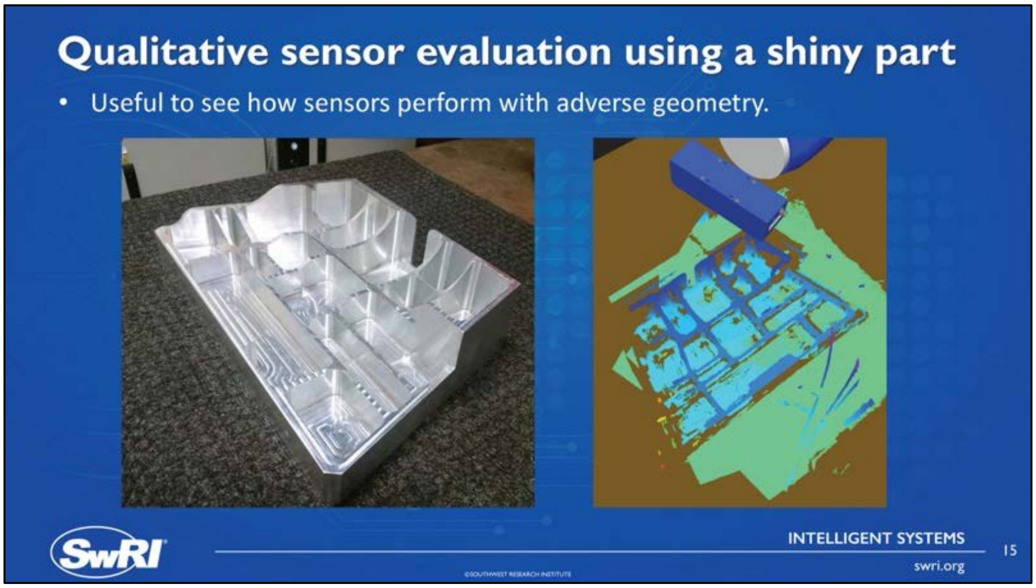


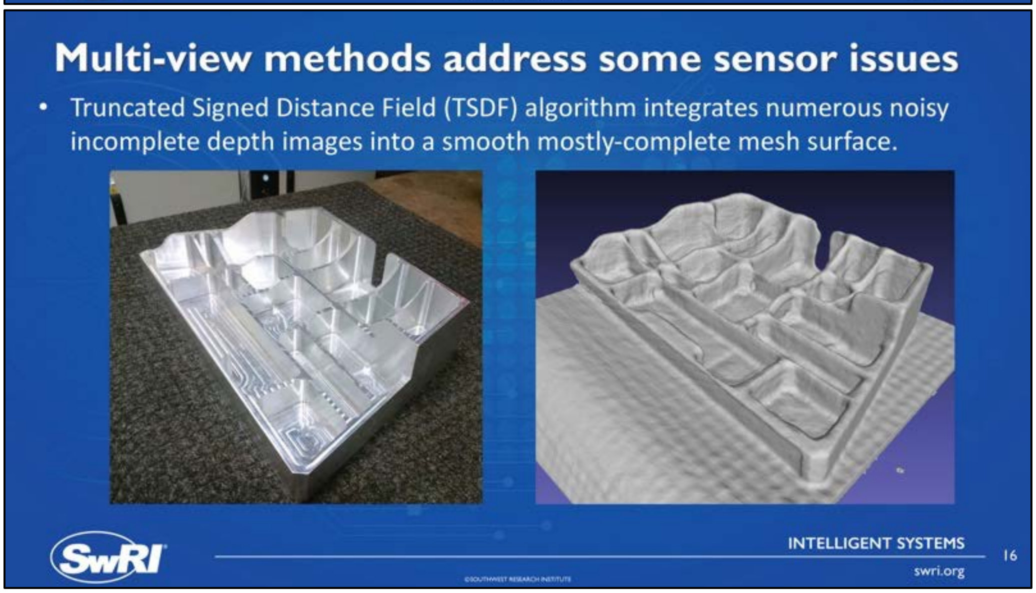


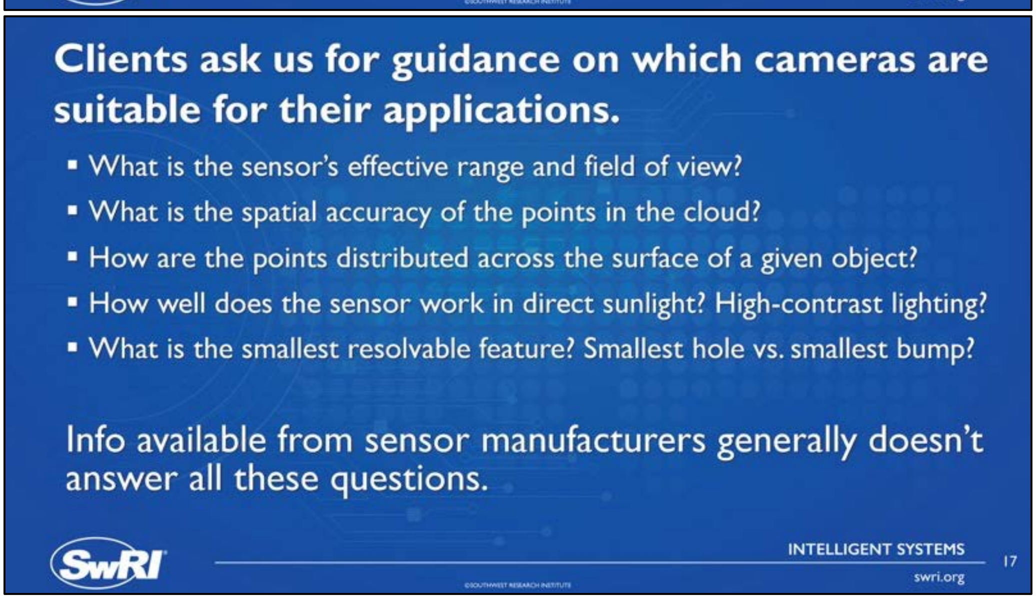
Comparing expected vs. actual corner positions shows asymmetric skew in point cloud






Qualitative sensor evaluation using a shiny part
• Useful to see how sensors perform with adverse geometry.
SwRI
INTELLIGENT SYSTEMS
swri.org
15


Multi-view methods address some sensor issues
• Truncated Signed Distance Field (TSDF) algorithm integrates numerous noisy incomplete depth images into a smooth mostly-complete mesh surface.
SwRI
INTELLIGENT SYSTEMS
swri.org
16


Clients ask us for guidance on which cameras are suitable for their applications.
▪ What is the sensor's effective range and field of view?
▪ What is the spatial accuracy of the points in the cloud?
▪ How are the points distributed across the surface of a given object?
▪ How well does the sensor work in direct sunlight? High-contrast lighting?
▪ What is the smallest resolvable feature? Smallest hole vs. smallest bump?
Info available from sensor manufacturers generally doesn't answer all these questions.
SwRI
INTELLIGENT SYSTEMS
swri.org
17

A structured light scanner: Photoneo PhoXi

Scanner model	XS	S	M	L	XL
Resolution (3D points)	3.2 M	3.2 M	3.2 M	3.2 M	3.2 M
Scanning range (mm)	161 - 205	384 - 520	458 - 1118	870 - 2150	1680 - 3780
Optimal scanning distance (mm)	181	442	650	1239	2326
Scanning area (mm)	118 x 78	360 x 286	590 x 421	1082 x 802	1954 x 1509
Point to point distance(mm)	0.057	0.174	0.286	0.524	0.947
Calibration accuracy (mm)	0.025	0.050	0.100	0.200	0.500
Temporal noise (mm)	0.030	0.050	0.100	0.190	0.400

SwRI INTELLIGENT SYSTEMS swri.org 18

©SOUTHWEST RESEARCH INSTITUTE

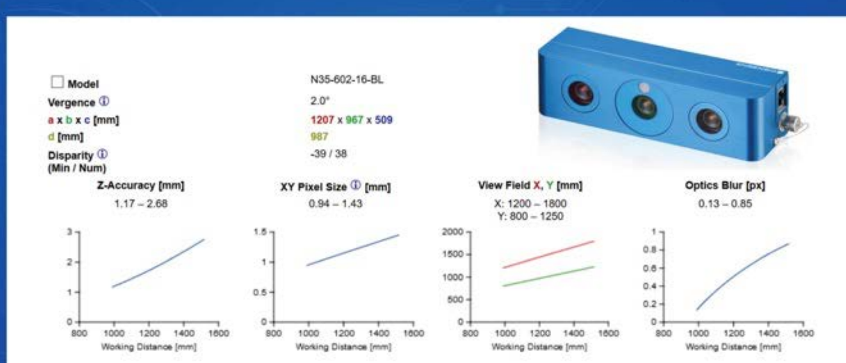
Would be useful to independently quantify each stage of 3D perception

- 2D feature detection algorithm
  - Errors in pixel positions due to inaccurate camera intrinsics or lens blur.
  - Metric for number of correspondence features per unit surface area?
- 3D point position calculation
  - Error due to inaccurate position/orientation between stereo cameras.
  - Important to separate theoretical optimal behavior from actual in-practice performance.

SwRI INTELLIGENT SYSTEMS swri.org 20

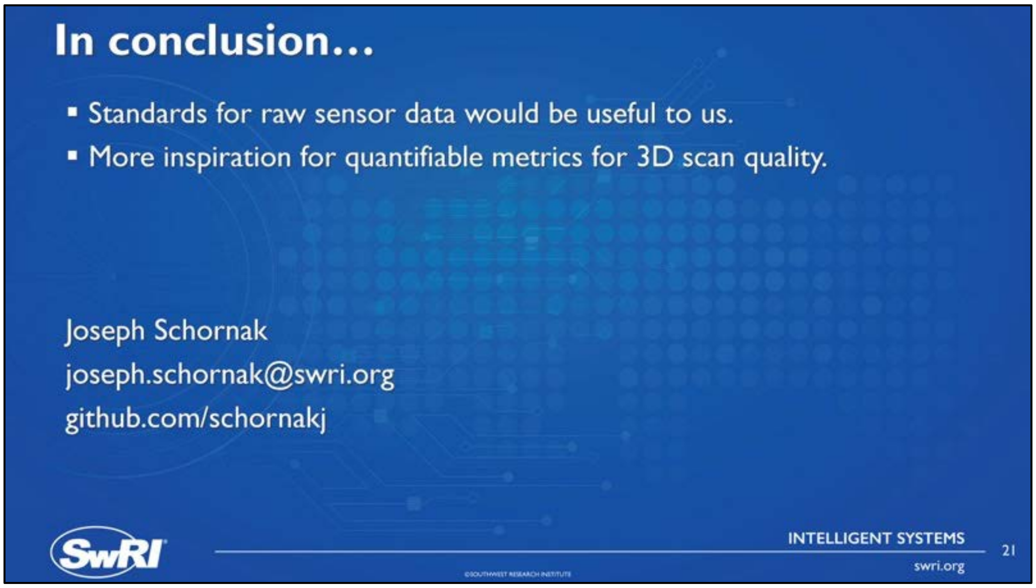
©SOUTHWEST RESEARCH INSTITUTE

A stereo camera: Ensenso N35



SwRI INTELLIGENT SYSTEMS swri.org 19

©SOUTHWEST RESEARCH INSTITUTE


In conclusion…
Standards for raw sensor data would be useful to us.
More inspiration for quantifiable metrics for 3D scan quality.
Joseph Schornak
joseph.schornak@swri.org
github.com/schornakj
SwRI
INTELLIGENT SYSTEMS
swri.org
21

B.7. Using 3D vision to control robots in dirty, industrial environments by Jared Glover, Capsen Robotics

We need ***smarter*** robots

$620 billion

- U.S. labor cost / year on "physical work done in highly structured and predictable environments" that **could be automated** with "currently demonstrated technology." *~ 2017 McKinsey study*

$2 billion

- 2018 U.S. robot sales

50 to 1

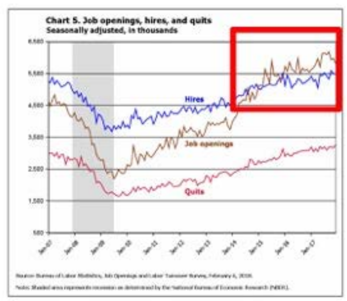
- Person-to-robot worker ratio in U.S. manufacturing (14 to 1 in S. Korea)

CapSen ROBOTICS

https://www.mckinsey.com/featured-insights/digital-disruption/harnessing-automation-for-a-future-that-works
https://ifr.org/ifr-press-releases/news/us-robot-density-now-more-than-double-that-of-china-ifr-says

We need smarter robots ***now***

- **Labor shortages due to:**
  - **3.5%** unemployment
  - **10,000** baby boomers reach retirement every day



- ***550,000*** *American manufacturing workers injured on the job in 2017* **(272 deaths)**

CapSen ROBOTICS

https://www.mckinsey.com/featured-insights/digital-disruption/harnessing-automation-for-a-future-that-works
https://injuryfacts.nsc.org/work/work-overview/work-safety-introduction/

Hard jobs for robots

These tasks are:

- Physically demanding
- Extremely repetitive

But still require a great deal of spatial intelligence:

- Fine manipulations
- Hand-eye coordination
- Picking in clutter
- Operating machinery

https://drive.google.com/open?id=14SuiUOtyq7SjQO8P6obsjFPjQ5XPeWFO

CapSen ROBOTICS

Our Solution

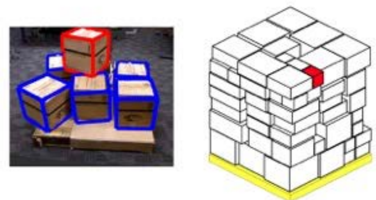
CapSen PiC™

3D vision software, full motion planning & control

- **Pick objects from clutter**
- **Precise placement** (position & orientation)
- **High-level task planning:**
  - Disentanglement
  - Stacking / Unstacking
- **Fast cycle time** (1 sec. image processing / planning)
- **Wide range of part sizes** (1 cm – 1 meter)
- **Works with any robot**
- **Supports multiple 3D cameras**

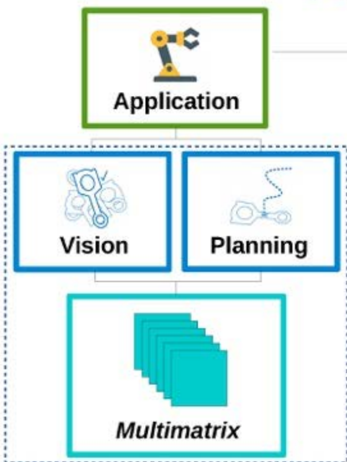
Bin Picking

Palletizing / Depalletizing



CapSen ROBOTICS

Our Platform



To meet your challenges, we offer:

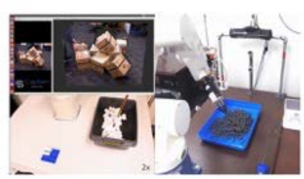
- **Existing solutions**
  - Bin picking
  - Machine tending
  - Palletizing / Depalletizing
  - 3D Scanning / Object Recognition
- **Advanced R&D capabilities**
  - Custom systems for your needs

CapSen SDK
Extensive software infrastructure
Developed in-house

GPU-accelerated
Multimatrix—a CapSen-developed software library for parallel matrix computations

CapSen ROBOTICS

Our demo videos



Overview of CapSen capabilities
https://drive.google.com/open?id=1qKbHFjBaSNHpiKN-H0xRsexRuaslxVQI

Hook picking robot we recently installed at a wire & spring factory:
https://youtu.be/fcvulzMQ1kg

Picking springs:
https://photos.app.goo.gl/HbnLtW4mSVq5aVVY9

Picking screws:
https://drive.google.com/a/capsenrobotics.com/file/d/10JxloRsfS1YCitcN6QZymAZ9H-pHWesn/view

Picking bolts with a magnet:
https://drive.google.com/file/d/1Hz4MOgHcRHC2ZFT5UHKwjWE8OFvog1Kc

Picking bolts with fingers:
https://drive.google.com/open?id=1yBsN0STPLhlWTPqYK_BOx4ZtU-LfUZTM

CapSen ROBOTICS

Recent installations

Wire & Spring Manufacturer

- **Pick** hooks out of a bin and insert them into a press

Fortune 500 Nuclear Tech. Company

- **Scan** and **spray** objects the robot has never seen before

CapSen ROBOTICS

Competitive advantage

- Complete solution (vision + **motion planning** + control)
- **World's first entangled-part bin picking**
- Can detect & pick **very small parts** (screws, nuts, etc.)
- Multiple competitors in **bin picking**
- Most provide only vision + grasp analysis (not full motion planning)

- We're not a bin picking company... ***we're a spatial intelligence company***

CapSen ROBOTICS

Application #1

- Traditional manufacturer (80+ years)
- Experts at what they do
- Family business

ace

ace wire spring & form co., inc.

CapSen ROBOTICS

- **Rich (Ace Wire Spring):**
  Nobody likes doing this job.
  Can we use a robot?

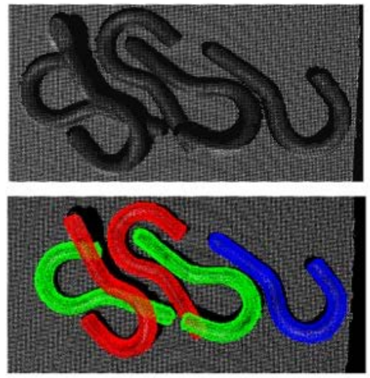
- **Rich (Ace Wire Spring):**
  Nobody likes doing this job.
  Can we use a robot?

- **Jared (CapSen):**
  - *Random bin picking*
  - *Small, complex parts*
  - *Precise placements*
  - *Parts can get entangled!*
  - ***Never been done before…***

- **Step 1**
  Scan in 3D models of the objects to be picked

- **Step 1**
  Scan in 3D models of the objects to be picked

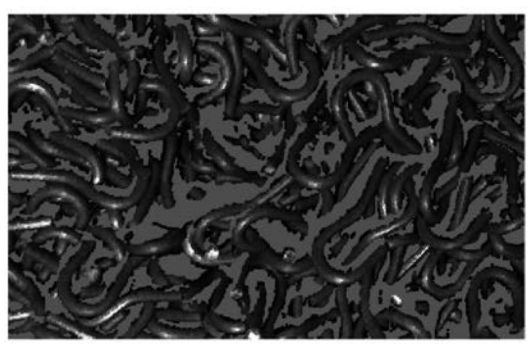
- **Step 2**
  Make sure our detection software can accurately locate the objects



- **Step 2...**
  Make sure our detection software can accurately locate the objects

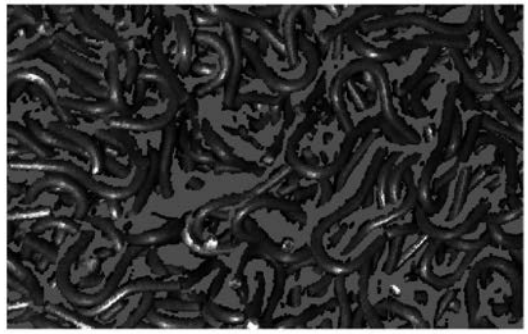
- **Step 2...**
  Make sure our detection software can accurately locate the objects

This is what the robot sees →



- **Step 2...**
  Make sure our detection software can accurately locate the objects

1. Select a 3D sensor
- PhoXi M

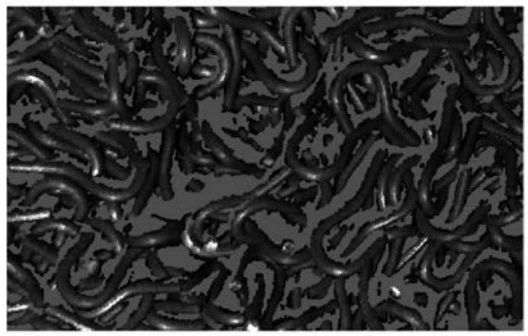


CapSen

- **Step 2...**
  Make sure our detection software can accurately locate the objects

1. Select a 3D sensor
- PhoXi M

2. Tune sensor params
- Exposure times
- Proj. patterns (number & orientation)
- Noise filtering



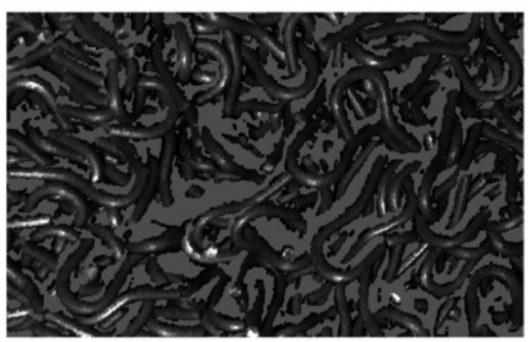
CapSen

- **Step 2...**
  Make sure our detection software can accurately locate the objects

1. Select a 3D sensor
- PhoXi M

2. Tune sensor params
- Exposure times
- Proj. patterns (number & orientation)
- Noise filtering

3. Tune CapSen params
- # detections vs. processing time
- (Optional) Train a new ML model



CapSen

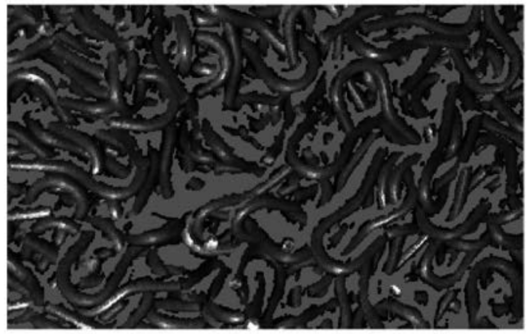
- **Step 2...**
  Make sure our detection software can accurately locate the objects

1. Select a 3D sensor
- PhoXi M

2. Tune sensor params
- Exposure times
- Proj. patterns (number & orientation)
- Noise filtering

3. Tune CapSen params
- # detections vs. processing time
- (Optional) Train a new ML model

4. Test
- If not good enough, return to step 1, 2, or 3



CapSen ROBOTICS

- **Step 3**
  Select a robot arm

PRECISE AUTOMATION

Repeatability: 30 microns

CapSen ROBOTICS

- **Step 3**
  Select a robot arm

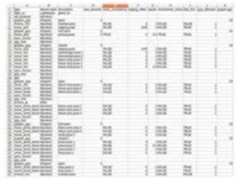
PRECISE AUTOMATION

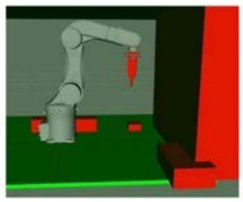
Repeatability: 30 microns

- **Step 4**
  Design fingers & fixtures

CapSen ROBOTICS

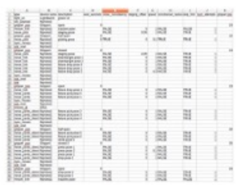
- **Step 5**
  Task configuration

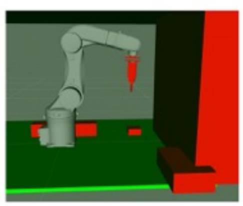




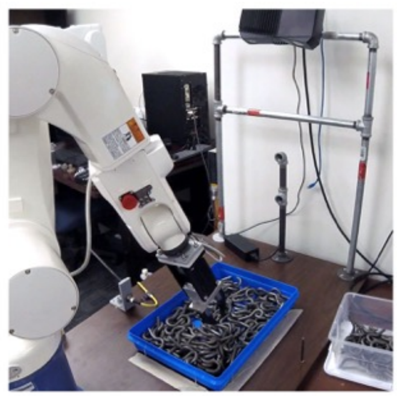
CapSen ROBOTICS

- **Step 5**
  Task configuration





- **Step 6**
  Full system test @CapSen



CapSen ROBOTICS

- **Step 7**
  Install @Ace

https://youtu.be/fcvulzMQ1kg

CapSen ROBOTICS

Application #2

Fortune 500 Nuclear Tech. Co.

- **Scan** and **spray** objects the robot has never seen before

CapSen ROBOTICS

Application #2

Fortune 500 Nuclear Tech. Co.

- **Scan** and **spray** objects the robot has never seen before

1. Some types of sensor noise are OK (multiple views are stitched together)
- Missing points
- Unbiased, Gaussian noise

CapSen ROBOTICS

Application #2

Fortune 500 Nuclear Tech. Co.

- **Scan** and **spray** objects the robot has never seen before

1. Some types of sensor noise are OK (multiple views are stitched together)
- Missing points
- Unbiased, Gaussian noise

2. Some types of 3D modeling errors are OK
- Small holes
- Missing details

CapSen ROBOTICS

Application #2

Fortune 500 Nuclear Tech. Co.

- **Scan** and **spray** objects the robot has never seen before

1. **Some types of sensor noise are OK (multiple views are stitched together)**
   - Missing points
   - Unbiased, Gaussian noise
2. **Some types of 3D modeling errors are OK**
   - Small holes
   - Missing details
3. **System failures:**
   - **Collision**
     (Due to missing object part)
   - **Incomplete coverage**
     (Due to incorrect model shape)

Lessons?

1. **Many 3D vision-guided robotic applications (including bin picking) are still at the bleeding edge**
   - Require customization / parameter tuning of sensors & software for each installation

Lessons?

1. **Many 3D vision-guided robotic applications (including bin picking) are still at the bleeding edge**
   - Require customization / parameter tuning of sensors & software for each installation
2. **Different applications can have vastly different performance metrics**
   - This argues for having many standards to choose from and/or complete end-to-end system testing

Lessons?

1. **Many 3D vision-guided robotic applications (including bin picking) are still at the bleeding edge**
   - Require customization / parameter tuning of sensors & software for each installation
2. **Different applications can have vastly different performance metrics**
   - This argues for having many standards to choose from and/or complete end-to-end system testing
3. **Performance metrics depend on many inputs**
   - Sensor, sensor params, software params, environment, etc...

CapSen Robotics

Lessons?

1. **Many 3D vision-guided robotic applications (including bin picking) are still at the bleeding edge**
   - Require customization / parameter tuning of sensors & software for each installation
2. **Different applications can have vastly different performance metrics**
   - This argues for having many standards to choose from and/or complete end-to-end system testing
3. **Performance metrics depend on many inputs**
   - Sensor, sensor params, software params, environment, etc...

CapSen Robotics

Appendix C: Individual Group Scores and Rankings

This appendix contains, for each group and for each idea, the scores for each element and the group ranking for an idea.

Table E: Groups 1, 2, and 3 Scores and Rankings

		SCORES														
IDEA	**DESCRIPTION**	**Crit.**			**Ach.**			**Scope**			**Effect**			**Ranking**		
	Group Number:	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
Ambient conditions	Standards for measuring the effects of changes in ambient conditions (lighting, temperature, humidity, vibrations, EMF interference, background specular reflections, etc.) on the 3D perception system's part-pose measurement performance.	3	3	1	2	1	2	2	1	1	3	3	3	3	2	2
Performance due to cluttered versus uncluttered scenes	Standards for measuring a 3D perception system's ability to measure the 6DOF pose of a single part presented alone vs. a part presented within a cluttered environment.	1	2	1	1	1	1	1	1	2	2	2	3	1	1	2
Performance due to occlusions	Standards for measuring the effects of part occlusion (self-occlusions or occlusions by other parts) on the 3D perception system's part-pose measurement performance.	1	2	1	1	1	1	2	1	2	1	2	3	1	1	2
Temperature stability	Standards for measuring the effects of changes in a 3D sensor's internal temperature on the 3D perception system's part-pose measurement performance.	1	1	1	1	1	1	2	3	3	1	1	2	1	1	2
Dynamic performance	Standards for measuring the effects of sensor (or object) motion on a 3D perception system's part-pose measurement performance (e.g., ASTM E3064).	2	1	3	2	1	2	1	3	2	3	1	3	2	1	3
Changes in performance throughout a perception system's field-of-view (FOV)	Standards for measuring a 3D perception system's performance throughout its FOV.	3	3	3	1	1	2	2	3	2	3	3	3	2	3	3
Measurement volume specification/verification	Standards for measuring a 3D perception system's measurement volume (FOV, measurement range, calibrated distance, standoff distance, etc.)	3	3	3	1	1	2	1	3	2	3	3	3	2	3	3
Interoperability	Standard protocols, data formats, or interfaces to allow sensors from different vendors to work with software/robots from different vendors.	3	2	3	3	1	2	1	2	3	3	2	3	3	2	3

Latency	Standards for measuring the time between when a perception system is commanded to take a measurement and when a usable measurement is available to other systems, with possible definitions for "integration time," "frame rate," and "real-time" (e.g., ASTM 3124-17).	3	2	3	2	1	2	2	2	2	2	2	3	2	2	3
Time synchronization	Standards for measuring the time synchronization between different 3D sensors or systems (e.g., IEEE 1588).	1	1	3	2	3	2	1	2	2	2	1	3	1	2	3
Cycle time	Standards for measuring the time for a robotic system to estimate the 6DOF pose of a part, grip the part, and deliver the part to its final destination. (E.g., "cycle time" could be defined as the time it takes between the command to the 3D perception system to measure the 6D pose of a part until the pose is available for the robot to use - or until the robot acquires the part).	1	3	3	1	1	1	1	3	1	2	3	3	1	3	2
XYZ linearity	Standards for measuring how linear a 3D perception system's measurements are in x, y, and z.	1	1	3	1	1	2	1	1	2	1	1	3	1	1	3
Pointcloud XYZ resolution	Standards for evaluating the smallest measurements that a system can achieve in the X, Y, and Z directions for 3D perception systems that produce pointclouds from a single sensor or multiple sensors.	3	3	3	1	1	2	1	2	2	3	3	3	2	2	3
Depth map XYZ resolution	Standards for evaluating the smallest measurements that a system can achieve in the X, Y, and Z directions for 3D perception systems that produce depth maps.	3	3	3	1	1	1	1	2	2	3	3	3	2	2	2
2D image XY resolution	Standards for evaluating the smallest measurements that a system can achieve in the X and Y directions for 3D perception systems that produce 2D images.	3	3	3	1	1	3	1	2	3	3	3	3	2	2	3
Part position resolution	Standards for evaluating the smallest changes of a part's position along the X, Y, and Z axes that a 3D perception system can measure.	3	3	3	1	2	1	1	2	2	3	3	3	2	3	2
Part orientation resolution	Standards for evaluating the smallest changes of a part's orientation about the X, Y, and Z axes that a 3D perception system can measure.	3	3	3	1	2	1	1	2	1	3	3	3	2	3	2
Functional safety	Standards for evaluating a 3D perception system's functional safety (i.e., its ability to properly handle likely human errors, hardware failures and operational/environmental stress - e.g., ISO 26262).	2	3	3	3	3	2	2	3	1	3	1	3	3	3	2
Eye safety over FOV	Standards for measuring the eye safety of a sensor's active illumination across its entire FOV.	3	3	3	3	3	2	2	3	2	3	1	3	3	3	3

Output quality	Standards for measuring a 3D perception system's ability to quantify the quality of the output (e.g., values for different types of errors, confidence in 6D pose, false positives, measurement dispersion over time, etc.).	3	3	3	1	1	1	1	1	1	3	3	3	2	2	2
Standard reference objects or artifacts	Develop standard reference objects that can be used for benchmarking and/or calibrating a 3D perception system's performance (e.g., interreflections, concave vs. convex parts, curved vs. planar surfaces, etc.).	3	3	2	2	2	3	1	2	3	3	3	3	2	3	3
Ability to resolve geometric features	Standards for measuring a 3D perception system's ability to resolve geometric features (e.g., edges and corners) on standard reference objects.	3	3	3	1	1	1	1	2	2	3	3	3	2	2	2
Error against traceable targets	Standards for using standard reference objects to evaluate a 3D perception system's errors.	3	3	2	1	2	2	2	2	3	3	3	3	2	3	3
Performance due to part material properties	Standards for measuring the effects of different part material properties on the 3D perception system's part-pose measurement performance (e.g., effects of light penetration).	2	3	2	1	1	1	2	1	1	2	3	3	2	2	2
Performance due to part surface properties	Standards for measuring the effects of different part surface properties on the 3D perception system's part-pose measurement performance (e.g., diffuse vs. specular reflections, reflectance, etc.).	2	3	2	1	1	1	2	1	1	2	3	3	2	2	2
Part reflectance	Standards for measuring part reflectance (e.g., parts with curved surfaces, multifaceted parts, parts with multiple reflectivities, etc.)	2	2	2	1	1	2	3	1	2	2	2	3	2	1	2
System-to-part suitability	Standards to determine whether a 3D perception system is appropriate for determining the pose of a part for a particular application. (E.g., is a particular system useful for small, metal automotive parts?)	1	1	3	1	2	1	1	1	1	1	1	2	1	1	2
Static performance	Standards for evaluating a 3D perception system's static part-pose measurement performance (e.g., ASTM E2919).	3	3	1	3	1	2	2	1	1	2	3	2	2	2	1
Depth error	Standards for evaluating a 3D perception system's depth error.	3	2	3	1	1	2	2	1	1	3	2	2	3	1	2
Bit precision resolution	Standards for measuring a 3D perception system's ability to define the precision of the data	1	1	1	1	1	2	1	1	3	1	1	2	1	1	2
Power requirements	Standards for measuring the power consumption of perception systems (e.g., spikes in power, startup power, etc.).	2	1	3	2	3	3	2	1	3	2	1	2	2	1	3

Frame rate	Standards for measuring/defining a perception system's actual frame rate.	3	1	2	2	1	2	3	1	3	2	1	2	2	1	2
Repeatability	Standards for measuring the variation of test results over a short and long periods of time.	2	3	3	1	1	2	3	1	1	2	3	3	2	2	2
Standard robot platform for complete system testing	Standard system setup for testing integration of new vision system; e.g. send robot tool center point to desired location from camera system to measure system level accuracy.	3	3	3	1	1	1	1	2	2	3	3	3	3	2	2
Reliability & Robustness	Standards for measuring performance throughout long-term use or exposure to regular work environmental conditions (e.g., vibration, temperature, etc.).	3	3	3	1	2	2	2	2	1	3	3	3	3	3	2
Power connector interface	Standards for sensor power connections to enable interchangeability of different 3D perception systems.	2	1	3	2	1	3	2	1	3	2	1	3	2	1	3
Data compression	Standards for 3D data compression to benefit data storage and transmission (e.g., .e57 format).	3	2	3	2	1	2	2	1	2	2	2	3	2	1	3
Computation power of host computer	Standards for evaluating the computation resources required to achieve certain latency of 3D perception systems that require off board processing.	3	2	2	1	1	1	1	1	3	2	2	2	2	1	2
Calibration quality	Standards for evaluating intrinsic and extrinsic camera calibration quality.	2	3	3	1	1	3	2	1	3	1	3	3	1	2	3