robotics business review

WHITEPAPER

Intel RealSense Brings 3D Vision to Robots

Machine vision enables robots to see and better interact with their environment. The eyes of the robot are the cameras, while the brain is the central processing unit that interprets all the captured data.

Machine vision is nothing new, of course, but 3D depth sensing in real time has long been a major challenge for the robotics industry. While laser, infrared, and radar offer accurate depth sensing, the technologies would drive the price of autonomous robots through the roof.

Building robots with human-like visual intelligence that can see, understand, interact with, and learn from their environment is a challenge many in the industry are working hard to solve.

Intel is one of those companies. Intel RealSense technology has quickly become another low-cost 3D sensing solution that is helping developers build robots that can see like humans and sense depth and track motion.

This report will detail how Intel RealSense is giving the power of 3D vision to next-generation robots. We will analyze how RealSense is getting smarter thanks to artificial intelligence (AI), profile robots currently using RealSense cameras and discuss where the technology still needs to improve.

TABLE OF CONTENTS

REALSENSE TIMING JUST RIGHT

Israeli-based PrimeSense once was the leader in 3D sensing technology — Microsoft Kinect, another popular sensing platform for robots, was based off PrimeSense. But that all changed when Apple acquired PrimeSense for $360 million in November 2013 and stopped selling the products.

Brian Gerkey, CEO of the Open Source Robotics Foundation (OSRF), said the discontinuation of PrimeSense created a major void in robotics circles.

"PrimeSense was great, and Microsoft Kinect was a fantastic robot sensor. Every robotics lab bought as many Kinects as possible," Gerkey said. "Intel getting into the sensor game with RealSense was really smart. It really filled a gap."

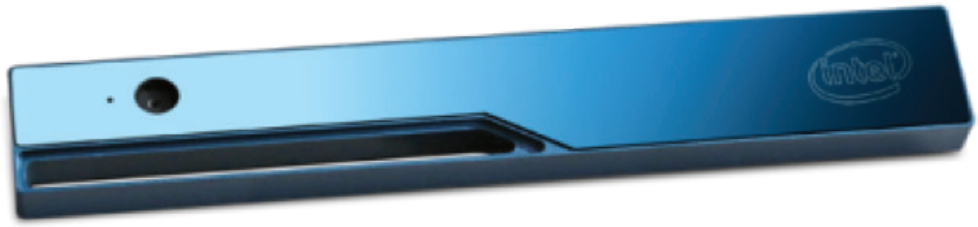
Prior to using RealSense, Savioke's autonomous mobile robot, Relay, used PrimeSense cameras to sense its environment. "PrimeSense was affordable, relatively available, and supported [the Robot Operating System] (ROS)," said Adrian Canoso, chief design officer, Savioke. "It was a really great sensor, but it started showing some age at the time of the Apple acquisition, as other sensors were coming to market. So the PrimeSense market was already starting to go away."

MEET THE INTEL REALSENSE LINEUP

INTEL REALSENSE D400 SERIES CAMERAS

Intel will soon release the RealSense D400 Series, which consists of the D415 and D435 Depth Cameras. Both cameras, which can be used for both indoor and outdoor applications, offer active stereo depth supporting 1280 x 720p resolution, 90 FPS dense depth stream and 2D RGB sensor for color. The cameras can be used across multiple platforms and support RealSense's open-source SDK. The chart below details the main differences between the two depth cameras.

INTEL REALSENSE DEPTH CAMERA D415	INTEL REALSENSE DEPTH CAMERA D435
FOV (horizontal x vertical x diagonal) 69° x 42° x 77°	FOV 91° x 65° x 100°
Dual rolling shutter sensors for up to 90 FPS depth stream	Dual global shutter sensors for up to 90 FPS depth stream



RealSense R200

INTEL REALSENSE CAMERA REAR R200

The R200 offers a long range, stereovision 3D imaging system. The small size of the R200 module provides flexibility to design into a wide range of products. The R200 camera is a USB 3.0 device that provides color, depth, and infrared (IR) video streams. The R200 consists of an IR laser projection system, two IR cameras (left and right) and one RGB camera. The depth video stream is generated with stereo vision technology assisted by the IR laser projector and the two IR imaging sensors.

Capable of being used for both indoor and outdoor applications, the R200 captures VGA resolution depth information at 60 frames per second (FPS) with more than 10 million depth point calculations per second and can be used both indoors and outdoors. It captures RGB data at 30 FPS.

Applications: scene capture, person tracking, object recognition and environment awareness

RealSense SR300

REALSENSE SR300

Intel RealSense Camera SR300: This USB-powered camera that offers a short range, coded light 3D imaging system. It is ideal for makers, educators, hardware prototyping and software development. The SR300 is designed for indoor applications and captures VGA resolution depth information at 60 FPS with more than 10 million depth point calculations per second. It captures RGB data at 30 FPS.

Applications: 3D scan, facial recognition and background segmentation

	INTEL REALSENSE R200	INTEL REALSENSE SR300	INTEL REALSENSE ZR300
Indoor Range	0.7m – 3.5m	0.2m – 1.5m	0.55m-2.8m
Outdoor Range	10m	N/A	>3.5m pending lighting conditions
Depth FoV (WxH)	56°x43°	68°x54°	*IR Laser Projector:* 60°x60° *Infrared Camera:* 59°x46° *Color Camera:* 68°x41.5° *Fisheye Camera:* 133°x100°
Depth FPS	30,60,90	30,60	30, 60
IR FPS	30,60,90	5-200	30, 60
Depth Method	Active Stereo IR	Coded Light IR	Active IR Stereo
Minimum F/W	1.0.72.04	3.21.0.0	N/A

INTEL REALSENSE CAMERA ZR300

RealSense ZR300

The ZR300 camera provides 6 degree of freedom motion, fisheye monochrome, color, depth, and IR video streams over a combination of USB and CSI interfaces. The low-powered ZR300 uses stereo vision to calculate depth, consisting of three cameras: a 1080p HD camera, an infrared camera, and an infrared laser projector.

The ZR300 integrates a depth camera for computing high-density depth (more than 10 million points per second) and a wide-field-of-view camera (VGA with >160-degree FOV) with an accelerometer-gyroscope combination for motion and feature tracking. Rounding out the six camera sensors on the device are an 8MP rear RGB camera and a 2MP front-facing RGB camera. The ZR300 provides depth data at VGA-resolution of 60 FPS and RGB data at 30 FPS.

Applications: SLAM, person tracking, object recognition, gesture recognition

ALL ABOUT THE DATA

Why is it important for robots to sense their environments? Well, the more data the robot captures, the more information the robot has to be more productive at the task at hand.

"There's been a big shift to data," said Sagi Ben Moshe, VP of New Technology Group, GM of RealSense. "Once you have the data, you can learn from it and create much smarter machine. Robots analyzing lots of data in real time is one of the trends changing the world."

Moshe called the RealSense cameras the eyes of the robot. "If you don't see the world, you can't accomplish your task. "You can't just give a robot a map and then it moves around a hotel. The robot needs to sense that something is coming and something is going. It needs real-time sensing to reach its destination. At 60 FPS, a robot can really understand the world in real time."

HOW REALSENSE HELPS ROBOTS

3D Depth Sensing: RealSense cameras can calculate the distance between objects, separating objects from the background layers. This gives much better object, facial and gesture recognition than a traditional camera, according to Intel.

Mapping and Navigation: RealSense SLAM uses a fisheye camera, accelerometer, gyroscope, and depth camera to track a system's movement in 6DoF. It also allows a location that was mapped previously to be recognized, which is known as re-localization. Tracking and re-localization allow robots to build and share knowledge about an environment.

Facial Recognition/Person Tracking: Identify faces in the camera's range or facial features on an individual face. Supports 78 landmark points for increased accuracy, true 3D face detection as well as roll, pitch, and yaw of the face. Can accurately detect and track up to four faces in a scene, at a range of up to 1.2 meters.

Obstacle avoidance: Whether it is a drone flying through the air or a robot navigating a home, RealSense can help robots identify and autonomously avoid objects. RealSense cameras can calculate the distance between objects and separate objects from the background layers behind them.

BENEFITS OF REALSENSE

Canoso said the perception story for robots is all about sensor fusion and that "overlapping sensors make them nimble in human environments." Canoso said the biggest benefit of RealSense is it uses stereo cameras with structured light to open up a lot of potential for operating robots indoors and outdoors.

There are many other benefits to using Intel RealSense for robotics applications, of course. But the sources *Robotics Business Review* interviewed, however, mentioned the same benefits: small form factor, low cost, support for the Robot Operating System (ROS), which is popular among robotics startups, and scalability.

Let's dive a little deeper into some of the aforementioned benefits:

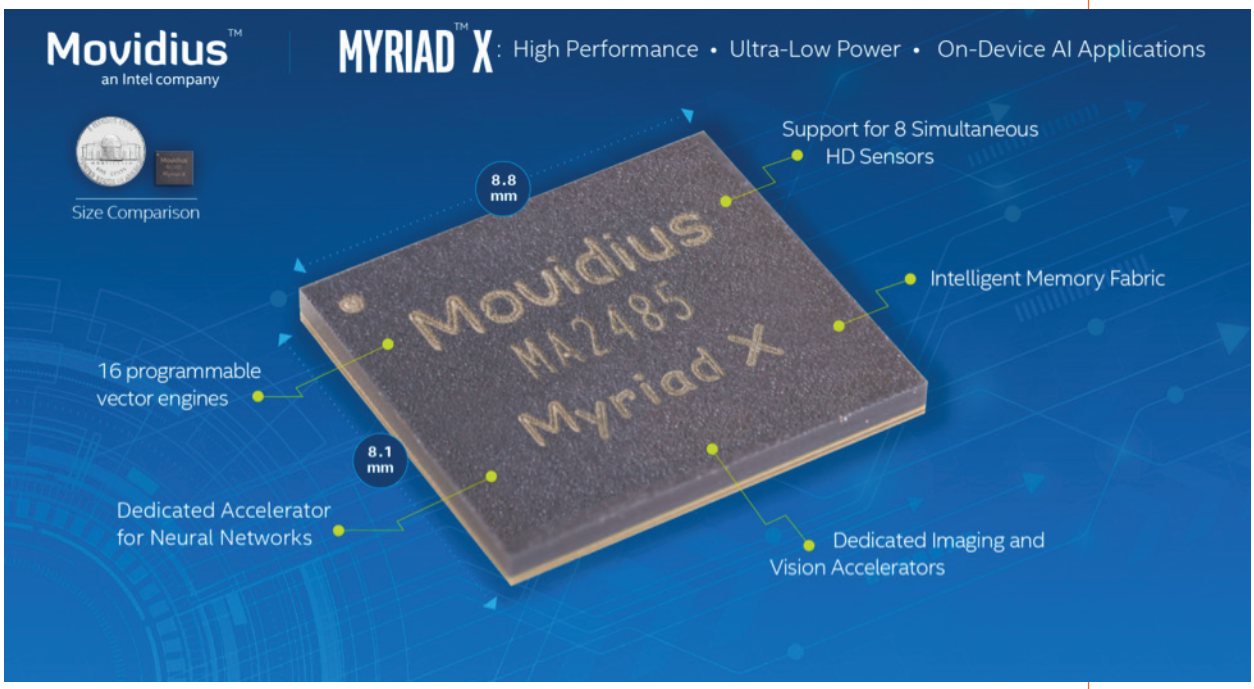
Small form Factor: Microsoft Kinect had a lot of upside as a 3D sensor, but Intel RealSense packs a more powerful punch in an even smaller package. Kinect was designed primarily for gaming, so size (24.9 cm x 6.6 cm x 6.7 cm) wasn't an issue. Size does matter when it comes to robotics, and the RealSense R200, for example, comes in at 101.6mm x 9.6mm x 3.8mm.

ROS Support: Much to the delight of robotics startups, Savioke launched a ROS wrapper for RealSense at the 2015 Intel Developer Forum. The wrapper allowed developers to better use the RealSense cameras. Savioke CEO Steve Cousins, of course, oversaw the development of ROS during his time running Willow Garage, the defunct robotics research lab.

"RealSense has really good APIs, especially the SLAM library," said Ricky Ye, founder and CEO of DFRobot. "With the Kinect, you needed a PC connected to it, which increased the power consumption. RealSense comes in a small footprint and allows a small robot to validate all the capabilities and greatly reduce overall cost of the whole platform."

Scalability: When Segway Robotics, which was purchased by Ninebot in 2015, started developing its Loomo robot (more on Loomo later), Ninebot VP of Robotics, Li Pu, chose RealSense in part because of Intel's ability to scale a product. "When we were exploring sensors, Intel was the only company we thought had the know-how and experience to scale up for mass production," Pu said.

Canoso agrees. "With Intel comes scalability," he said. "Intel sees the opportunity mobile robots provide. This domain is typically challenged by initial low volumes, but this is a several billion dollar industry. It takes a big company to nurture a young space and then bring it to mass."



Movidius Myriad X VPU (Credit: Intel Corporation)

REALSENSE GETS SMARTER WITH AI

In November 2016, Intel acquired Movidius, the computer vision and machine learning startup behind Google's Project Tango 3D sensing and mapping technology. With the acquisition, Intel said it gained a low-power, high-performance SoC platform for accelerating and embedding computer vision capabilities. And keeping power usage low is key to bringing RealSense to robots.

Less than a year after the acquisition, Intel introduced the Movidius Myriad X vision processing unit (VPU), which Intel said is the first VPU to ship with a dedicated Neural Compute Engine. Myriad X is a computer vision system on a chip (SOC) specifically designed to run deep neural networks at high speed and low power. It's ideally suited for enabling visual intelligence and autonomous capabilities in compact, battery-powered devices.

Intel said the key to the third-generation VPU is that it enables visual intelligence directly onboard the device instead of in the Cloud, which removes the time, power and bandwidth constraints of communicating with the cloud.

Capable of delivering more than four trillion operations per second (TOPS)[1] of total compute performance, according to Intel, Myriad X combines imaging, visual processing and deep learning in real time with:

Programmable 128-bit VLIW Vector Processors: Run multiple imaging and vision application pipelines simultaneously with the flexibility of 16 vector processors optimized for computer vision workloads.

"Overall performance is the architectural calculation based on maximum performance of operations-per-second over all available compute units. Application performance varies based on the application."

Savioke Relay Robot

Increased Configurable MIPI Lanes: Connect up to 8 HD resolution RGB cameras directly to Myriad X with its 16 MIPI lanes included in its rich set of interfaces, to support up to 700 million pixels per second of image signal processing throughput.

Enhanced Vision Accelerators: Use over 20 hardware accelerators to perform tasks such as optical flow and stereo depth without introducing additional compute overhead.

2.5 MB of Homogenous On-Chip Memory: The centralized on-chip memory architecture allows for up to 450 GB per second of internal bandwidth, minimizing latency and reducing power consumption by minimizing off-chip data transfer.

At press time, Intel had not released pricing of the Myriad X or detailed any of the first robotics-related applications for the system.

CASE STUDIES

The goal of Intel RealSense is to equip autonomous machines with human-like 3D perception to understand, interact with, and learn from their environment. Here is a look at how five robots are using RealSense to their advantage.

1. SAVIOKE RELAY ROBOT

Savioke's Relay is, perhaps, the most commercially successful robot to date that uses Intel RealSense. Savioke has made its name delivering items to

hotel guests, having completed 150,000-plus deliveries to date without any incidents, according to Canoso. Most of those deliveries have occurred in hotels, but the RBR50 company recently branched out into luxury apartment complexes and indoor logistics with FedEx and Kinkos.

All those deliveries were made possible, thanks in large part, to the Intel RealSense cameras Savioke used. The autonomous mobile robots were using the R200 camera, but Canoso said the robots will be updated with Intel's new D400 series.

Thanks to RealSense and an Intel Core i7 Processor, Relay process what it sees in real-time, recognizing people and objects as obstacles and either stops or moves around them. This is all part of 16 overlapping sensors that Canoso said makes Relay redundant and safe.

RealSense has played an important role in giving Savioke a 3D point cloud to analyze obstacle and terrain data for its Relay robots. "If you put a toilet paper roll over your eye, that's how robots see. So you need a lot of other sensors to help it get to where it's trying to go," Canoso said. "But Intel helps Relay thread the needle between people and objects, and it's really awesome to see something you've only seen in the movies come to life."

2. SEGWAY ROBOTICS LOOMO ROBOT

Singapore's 2017 MIT Hacking Medicine Robotics Hackathon is a perfect example of how improved sensing is unlocking new robotics applications. Twelve teams participated in the hackathon to turn Segway Robotics' Loomo robot, introduced at CES 2016, into an assistive robot for the elderly. Loomo uses the Intel RealSense ZR300 3D Camera to perceive its environment, including depth-sensing, facial recognition, person tracking, and mapping and SLAM navigation.

Team Botler, which won the $5,000 first-place prize, used the ZR300, and other sensors, to turn Loomo into a self-driving wheelchair. Using a custom electromagnetic coupling, Loomo was attached to an ordinary wheelchair and pulled the wheelchair around. Using the ZR300's facial recognition, Loomo can identify patients at a nursing home, for example, connect to their wheelchair, and pull them from one location to another.

"Perceptual capabilities are very important for the advancement of robotics, and vision is the key factor that differentiates Loomo from other products," said Pu. "Facial recognition and person tracking give Loomo more natural

Team Botler's Loomo Robot

interaction with users. Combining the RealSense's depth and RGB image data, two very distinct measurements, help Loomo achieve better recognition and accuracy."

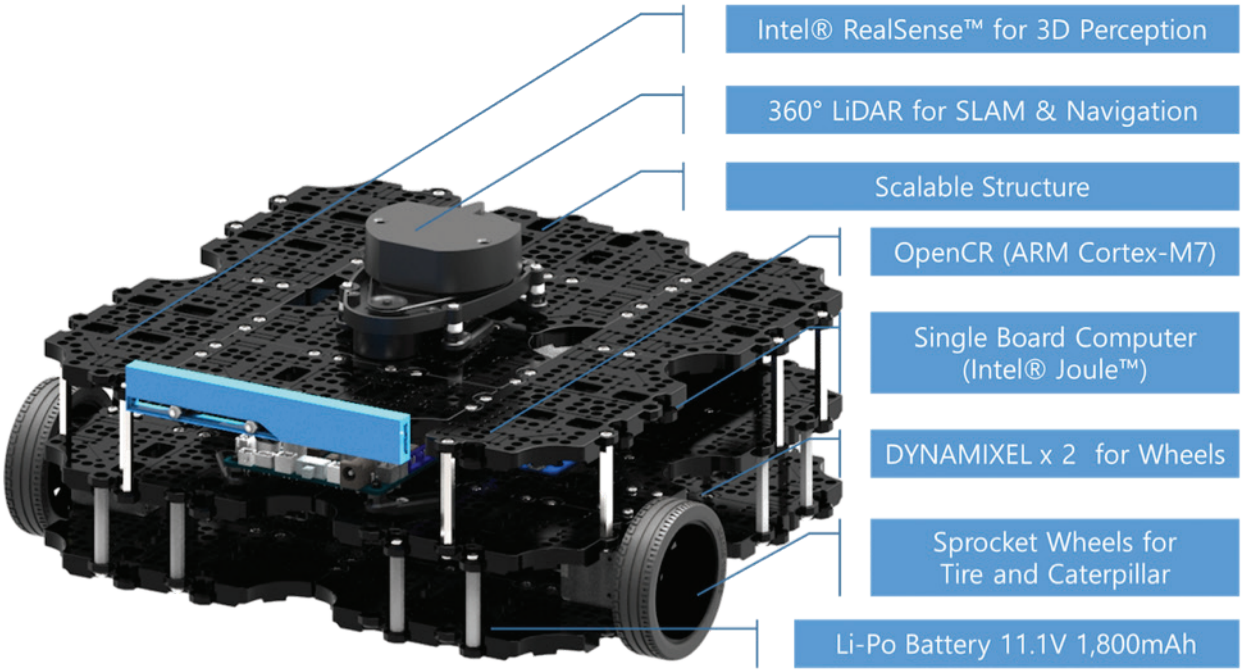
Sarah Zhang, senior director of robotics business operations at Segway Robotics, said Team Botler's solution frees up caregivers, who spend 10 percent of their time pushing wheelchairs around. "Loomo is very stable platform with many useful sensors, including Intel RealSense. And since Loomo can move, that helps elderly who lost the ability to move."

Other anticipated applications for Loomo, which recently entered mass production, include autonomous delivery, customer assistance in stores and hotels, and personal transportation.

3. NEXT-GEN TURTLEBOTS

TurtleBot, the popular open-source robotics R&D platform created at Willow Garage in 2010, is designed to bring complex technologies, such as autonomous navigation and robotic manipulation, into the hands of innovators and developers in a more approachable format. The next-generation models, the TurtleBot 2i and TurtleBot 3 Waffle, use Intel RealSense for advanced functionality.

The TurtleBot 2i uses the short-range Intel RealSense ZR300 SR300 that enables the MK3 robot arm to manipulate small objects, buttons, and tools in



TurtleBot 3 Waffle

its environment. Along with 360-degree LiDAR, the TurtleBot3 Waffle uses an Intel RealSense R200 to enable autonomous navigation. The TurtleBot 3 Waffle can also perform gesture recognition, object recognition and scene recognition based on 3D depth data captured by RealSense.

"The TurtleBot 2i is one of the cheaper mobile manipulator robots to ever hit the market," said former Interbotix Labs' principal engineer Andrew Dresner. "We rebuilt the hardware on the TurtleBot 2i from the ground up, added a robotic arm as a standard option, with Intel Joule and Intel RealSense. This is an industry first at the price level."

4. ASUS ZENBO ROBOT

Social robots are all the rage. Expectations are high, but the ability for companies to deliver these robots has been low. ASUS is one of the few companies that has delivered their social robot, launching Zenbo in Taiwan earlier in 2017, with plans to launch in China and the United States in the not-too-distant future.

Zenbo is a two-wheel robot assistant for your home that aims to provide assistance, entertainment, and companionship to people by moving about the home responding to voice commands, taking photos and videos, playing music, acting as a mobile security system and more.

To do all this, of course, Zenbo needs to sense its environment. Zenbo uses

ASUS Zenbo

to detect obstacles and navigate the environment, track body movements, and create an interactive experience with its users.

5. YUNEEC TYPHOON H DRONE

Intel is also working with leading drone companies, including Yuneec and its award-winning Typhoon H Drone. The Typhoon H features the Intel RealSense R200 and Intel Atom processor to map and learn its environment in 3D and autonomously detect and avoid obstacles by calculating the distance between itself and the object.

The R200 integrates with the Typhoon H's "Follow Me" mode to avoid objects while filming in any direction. This partnership marked the first time Intel RealSense was built into a drone. At the time the partnership was announced, Yuneec said "this is a major advancement over ultra-sonic collision prevention, which automatically stops Typhoon H short of obstacles, but cannot model the environment or intelligently re-route around obstacles."

CONCLUSION

While the ability for robots to sense and perceive their environment has come a long way, we are still in the infancy stage, according to the experts we interviewed. "There is high demand for a wider field of view and better dynamic range," said Ye. "That will help robots see better in strong light or dark

Yuneec Typhoon H Drone

environments. And on the computational side of the algorithms, when we have better data and higher resolution data, we'll need more processing power to run the algorithms."

Movidius and Intel RealSense both set out to give machines the power of sight. And on the surface it appears Myriad X and RealSense complement each other well. But only time will tell what impact Myriad X and RealSense will have on the next generation of intelligent machines that can track, navigate, and interact with their environment.