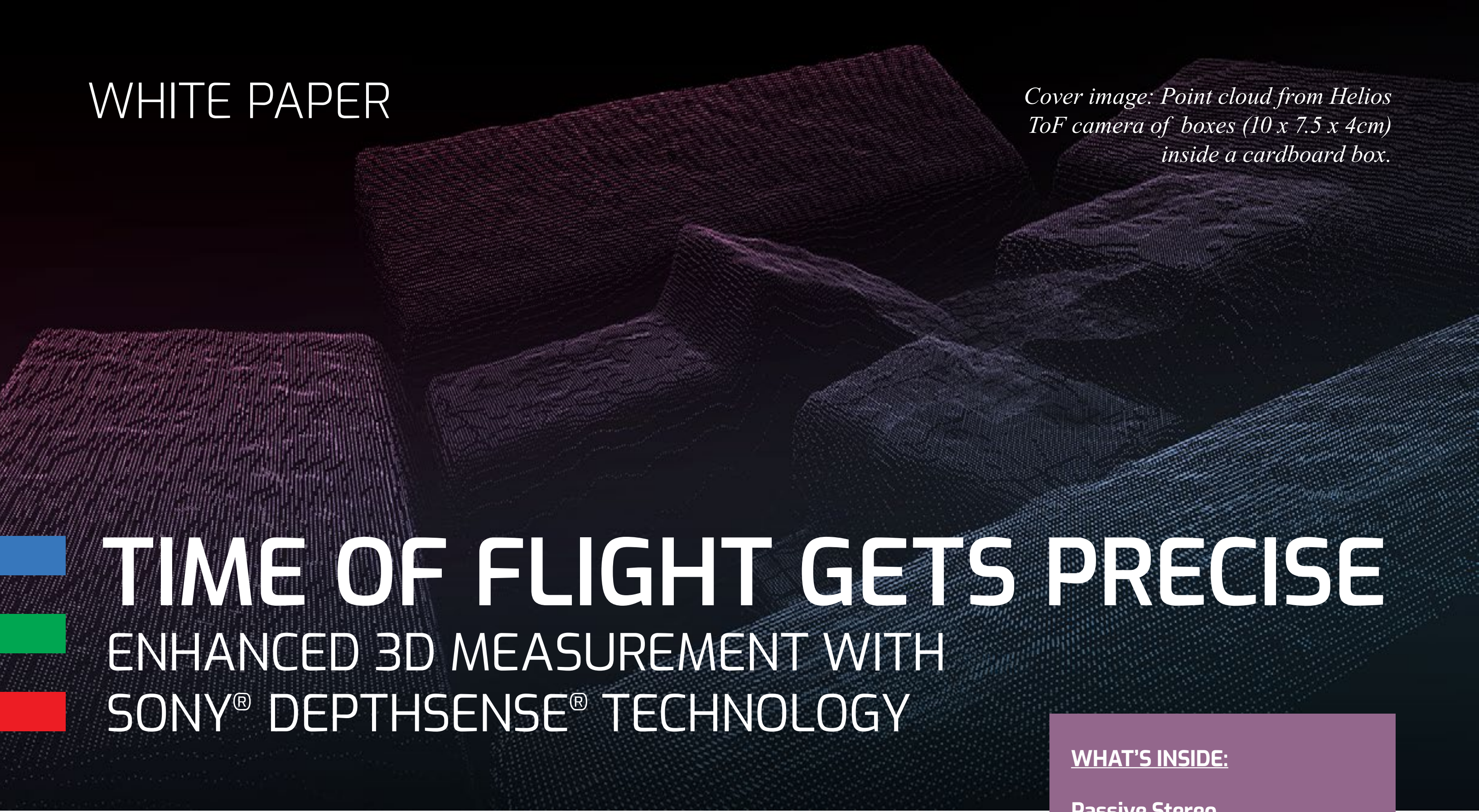


Cover image: Point cloud from Helios ToF camera of boxes (10 x 7.5 x 4cm) inside a cardboard box.

WHITE PAPER

TIME OF FLIGHT GETS PRECISE

ENHANCED 3D MEASUREMENT WITH SONY® DEPTHSENSE® TECHNOLOGY

WHAT'S INSIDE:

- **Passive Stereo**
- **Structured Light**
- **Time of Flight**
- **CW Phase-Shift Time of Flight**
- **Consider the Environment**
- **Lasting Benefits**

Three-dimensional (3D) imaging is used in many different industries ranging from both industrial pick and place, palletization/depalletization, warehouse, robotics and metrology applications to consumer-based products such as drones, safety and security, and patient monitoring applications. No one specific type of 3D technology can solve all these different types of applications and the features of each must be compared as to their suitability for each application ***(table 1)***.

3D imaging systems can be classified into both passive and active systems. In passive systems, ambient or broad fixed illumination is used to illuminate the object. Alternatively, active systems use various methods of spatially or temporally modulating light including laser line scanning, speckle projection, fringe pattern projection, or time of flight (ToF). **In both passive and active 3D imaging systems, reflected light from the object being illuminated is captured,** often by a CMOS-based camera, to generate a depth map of the structure and then, if required, a 3D model of the object.

In 3D imaging systems, captured data are represented in x, y and z coordinates known as a point cloud that contains horizontal (x), vertical (y) and distance information ***(figure 1)***. This is a collection of 3D points. These points can be represented as a 2D pseudo-colored depth map image where different colors represent the distance the object is from the 3D imager ***(figure 1)***. Capturing an image using a VGA sensor, for example, can produce a depth map of up to 640x480 pixels depending on the scanning method. Precision of the distance at each pixel is determined by its confidence value, which expresses the likelihood of the 3D measurement being correct.

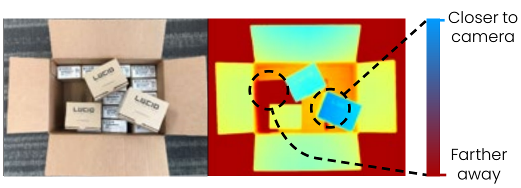


figure 1: Original image of the boxes and the depthmap (left). 3D point cloud image of the boxes (right)

	STEREO VISION (PASSIVE)	STRUCTURED LIGHT	TIME-OF-FLIGHT
Working Distance	Limited by baseline	Limited by baseline	Scalable with light source
Depth Accuracy	Low	High	Medium
Depth Map Resolution	Limited by texture of scene	Limited by light pattern	Full resolution
In-field Calibration Needed	Sometimes	Sometimes	No
Size	Increases with working distance	Increases with working distance	Compact
Cost	Low	High	Medium

table 1: Depending on your application, different 3D technologies will have certain advantages over others.

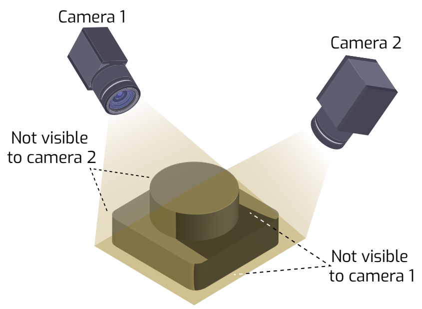


figure 2: Occlusion occurs when a part of an object cannot be imaged by one of the cameras.

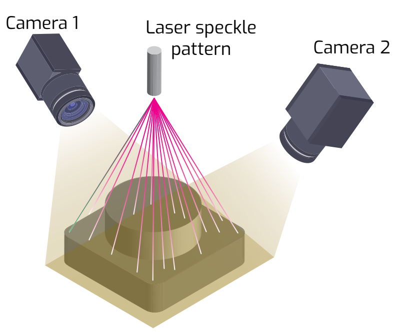


figure 3: Active stereo example using laser speckle pattern to create texture on object.

Passive Stereo

The standard passive stereo vision system uses two cameras located a fixed distance apart to capture a scene from two different positions. Using triangulation, depth information is extracted by matching features in both images. The closer an object is, the more its features will be shifted laterally apart in the two images. The magnitude of this shift, called the disparity, is used to calculate the distance of the matched features.

A 3D triangulation-based system can also be implemented using a single camera. Instead of using pairs of fixed stereo cameras, a single camera, often mounted on a robot, can position the camera at different positions around an object. By mapping multiple images using feature-extraction algorithms, 3D images can be re-created using specialized calibration techniques. The system accuracy will be limited to the positional accuracy of the robot.

Passive 3D systems may suffer from occlusion, where one part of the object or objects in the scene cannot be seen by one of the cameras. In this case, features within each captured image cannot be matched and the depth cannot be calculated. By using more than two cameras, occlusion effects that occur can in some cases be reduced ***(figure 2)***.

However, it is not only occlusion that presents a potential problem when determining depth information from stereo-based 3D systems. If an object has few features or faint features, for example, a completely white wall, then no matching correspondence pairs will be found between the two images. Without matching features in each image, it is impossible for depth information to be calculated.

Structured Light

To overcome the problem of finding features on nearly featureless objects, structured light systems project patterned light onto the objects. **Instead of searching for features that may be hard to see or even non-existent, the cameras only need to locate the well-defined patterns of light created by a light projector.** Structured light systems typically fall into one of two broad categories: active stereo systems or calibrated projector systems.

Active stereo systems operate in a very similar manner to passive stereo systems with the exception that artificial texture is projected onto the objects. The projected texture can be created by various means such as conventional reticle projectors with LED backlight, lasers with diffractive optic patterns, or laser speckle generated from diffusers ***(figure 3)***. The projected pattern is captured by a stereo camera system and feature matching is obtained by digital image correlation in the same way as passive stereo. Measurement accuracy will be limited by the inherent accuracy of the stereo camera in addition to the resolution of the projected pattern.

The second category of structured light systems discussed here makes use of calibrated projected patterns. Instead of just projecting texture for stereo correlation, the calibrated projector's pattern is accurately known and forms an integral part of the 3D measurement. Only a single camera is needed to compute depth from triangulation since the projector forms a known vertex on the triangle, analogous to the function of the second camera in a stereo camera pair. The simplest of these systems is the laser line scanner which projects a line of laser light across the surface of the object

(figure 4). To the camera, which is located off-axis from the laser, the reflected laser line appears distorted by an amount dependent on the distance to the surface. Scanning an object requires moving the laser line over the entire object and capturing a sequence of images. By analyzing these reflected line images, a geometric reconstruction of the surface shape of the object can be computed. Systems integrators can choose to build such systems using individual cameras and lasers that can be configured in different ways depending on the nature of the part to be imaged.

Passive stereo as well as structure light systems require calibration to produce meaningful results. Many manufacturers have introduced pre-calibrated structured light systems that embed both an illumination source and a camera into a single unit. An example of such a system is the Intel RealSense, providing a low cost, pre-calibrated structured light solution for application such as facial recognition. In this way, developers can deploy such systems more cost effectively. In industrial segments, structured light with one or more cameras is typically used in higher-end metrology applications since the accuracy can be much higher than passive stereo vision systems.

Time of Flight

While surface height resolutions of better than 10µm are achievable using laser scanners at short working distances, other applications demand longer range. For example, applications such as navigation, people monitoring, obstacle avoidance, and mobile robots require working distances of several meters ***(figure 6)***. In such applications, it is often simply necessary to detect if an object is present and measure its position to within a few centimeters.

Other applications such as automated materials handling systems, operate at moderate distances of 1-3m and require more accurate measurements of about 1-5mm. For such applications Time-of-Flight (ToF) imaging can be a competitive solution. **ToF systems operate by measuring the time it takes for light emitted from the device to reflect off objects in the scene and return to the sensor for each point of the image *(figure 7)*.**

ToF cameras use one of two techniques, pulse-modulation or continuous wave (CW) modulation. Pulse-modulation, also known as Direct ToF, involves emitting a short pulse of light and measuring the time it takes to return to the camera ***(figure 8)***. CW modulation emits a continuous signal and calculates the phase-shift between the emitted and returning light waves, which is proportional to the distance to the object ***(figure 9)***.

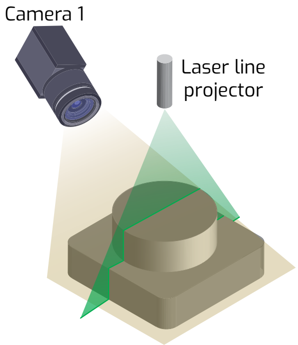


figure 4: Example of laser line scanner. Depth data is calculated by the amount of laser line distortion viewed by the camera.

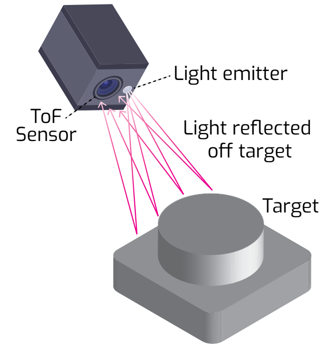
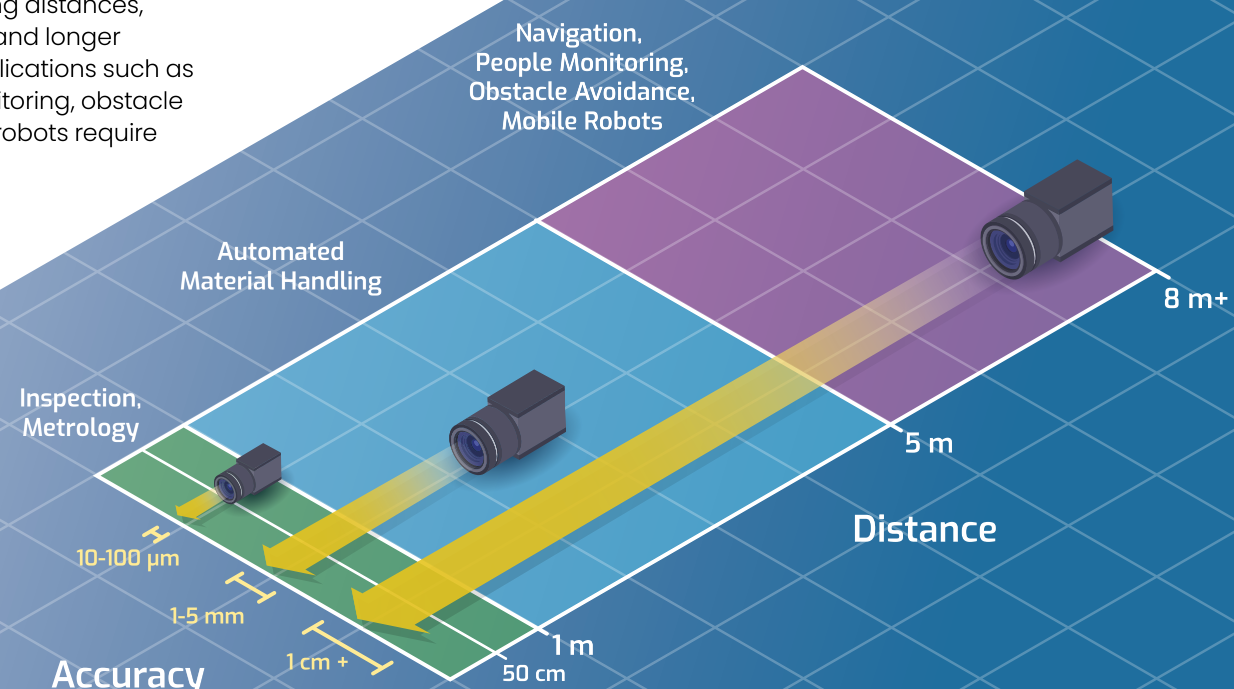


figure 7: ToF systems have a light source and a sensor.

figure 6: Different applications require different working distances. ToF technology is best suited for distances of 3m or less.



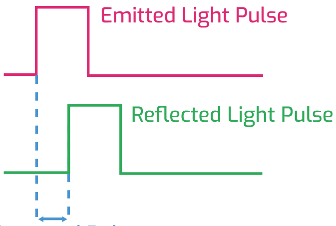


figure 8: Pulse-modulation emits a pulse of light and measures the time it takes to return. Because we know the speed of light, the time delay is proportional to the distance of the reflection surface.

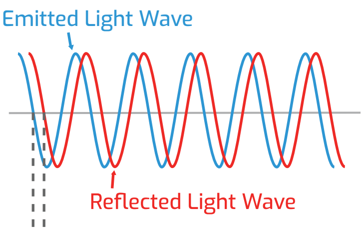
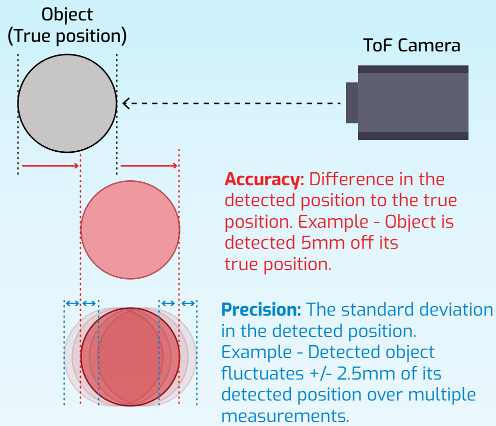


figure 9: Continuous wave-modulation technique measures the phase-shift between the emitted and reflected wave. The phase-shift is proportional to the distance of the reflection surface.

Important Terms: Accuracy and Precision

When talking about performance characteristics of ToF cameras it is important to differentiate between the terms "Accuracy" and "Precision".



CW Phase-Shift Time of Flight

CW Phase-shift based devices are available from several companies including Texas Instruments and Microsoft, with one of the newest sensors from Sony Semiconductor Solutions (Tokyo, Japan). Originally developed by SoftKinetics (Gewest, Belgium), now owned by Sony. The SoftKinetics technology features a Current Assisted Photonics Demodulation (CAPD) pixel structure capable of high-speed sampling with high efficiency ***(figure 10a)***. This ToF pixel technology is combined with Sony's backside illuminated sensor (BSI) technology to create the new DepthSense® ToF sensor. The BSI technology provides better light collection efficiency in NIR wavelengths ***(figure 10b)***. The new Sony IMX556PLR is a 1/2" sensor with 640x480 resolution, 10μm x 10μm pixels that runs at 60fps.

Recently, Sony's DepthSense sensor has been incorporated into the Helios™ ToF 3D camera from LUCID Vision Labs (Richmond, B.C., Canada). The camera can be operated at three working distances using light from four on-board 850nm VCSEL laser diodes modulated at different frequencies ***(figure 11)***. The camera has a 65° x 46° field of view and an operating range of 0.3 to 6m. At its

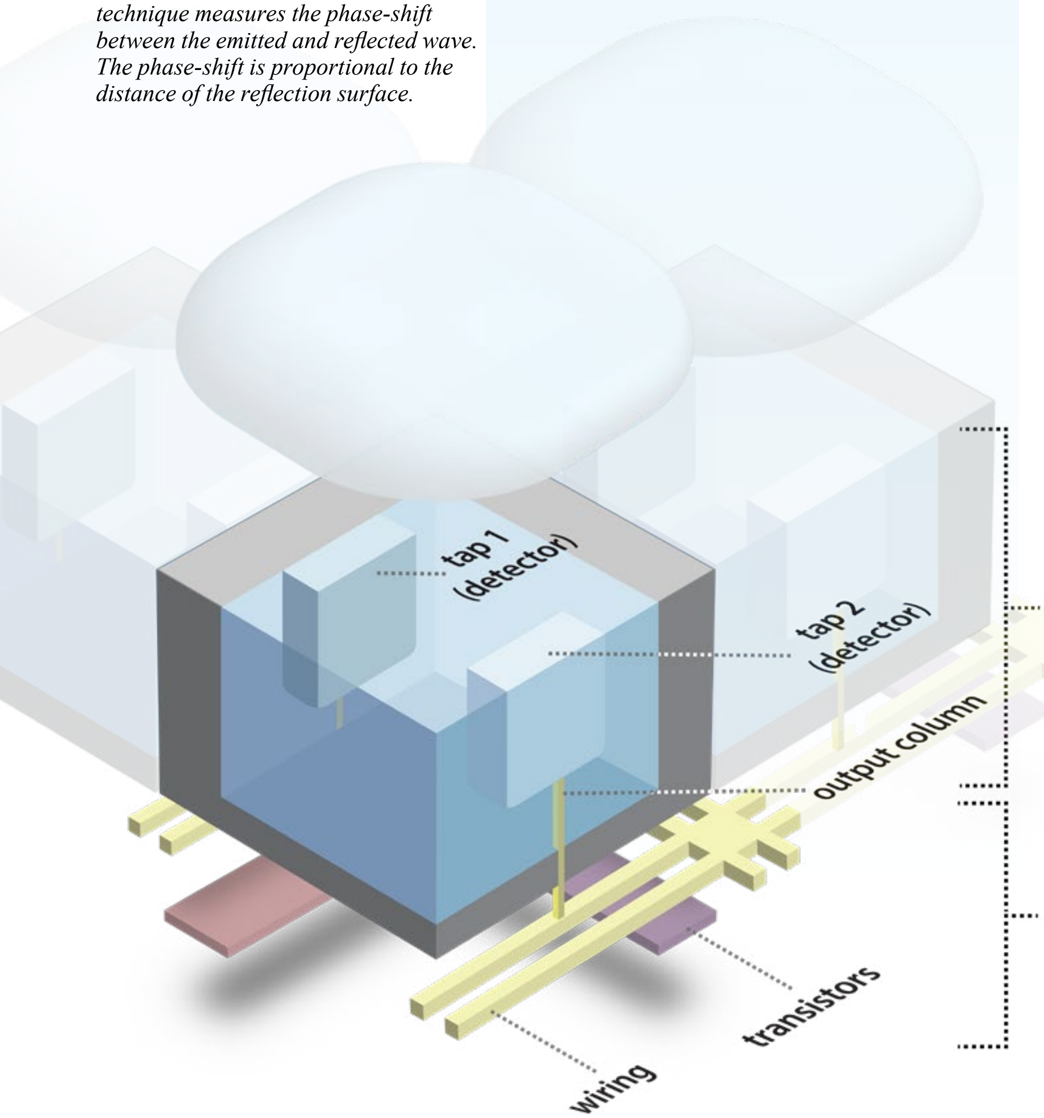


figure 10a: Sony's Current Assisted Photonics Demodulation (CAPD) pixel. Capable of high-speed modulation using drift current in the pixel, allowing for improved accuracy of distance measurement.

figure 10b: Sony's backside illuminated sensor (BSI) technology places the circuitry layer below the photodiode, allowing for increased sensitivity of the 850nm wavelength emitted by the VCSEL laser diodes.

highest modulation frequency of 100MHz the camera has better than 2.5mm precision and 5mm accuracy over its 0.3 to 1.5m working range. **The camera performs on-board processing producing 3D point cloud data** that can be read directly from the device instead of having to be calculated off-chip. The point cloud data can be transferred over the camera's GigE interface and further processed using LUCID's Arena software development kit (SDK), which allows users to write custom programs interfacing with the camera in C, C++ or C#.

In the past, camera vendors needed to supply drivers to their customers so that they could properly be configured and controlled by a host computer. With the introduction of GenICam, a generic programming interface for machine vision cameras, the camera interface is decoupled from the user's API, alleviating the need to write multiple camera drivers. Now, with the release of GenICam 3.0, depth and amplitude data or processed point cloud data can be read from cameras without any further data conversion required. Since the Helios camera is GenICam compatible, it can be very quickly integrated with off-the-shelf GenICam software packages that are currently available.

Consider the Environment

No 3D imaging technology meets the needs of every application. When choosing an active or passive imaging system several factors must be considered such as whether any additional light sources are present, what type of surfaces need to be imaged and whether any objects in the field of view are specular or will produce multiple reflections.

Light sources such as sunlight or specular reflections off shiny objects can easily saturate cameras used in stereo vision systems. Conversely, low light levels can produce noisy results ***(figure 12)***. For low lighting environments, slightly more expensive active stereo vision methods can be used. Applications such as metrology that require micron level precision measurements may need more expensive active laser systems placed relatively close to the object while other active laser systems such as scanning LIDAR can achieve much larger distances with lower absolute accuracy. For indoor environments with 0.5 to 6 m distance ranges, such as warehouses and factories, ToF technology provides a compact solution that can overcome certain lighting and object texture challenges where other 3D technologies could not.

figure 11: Helios imager board

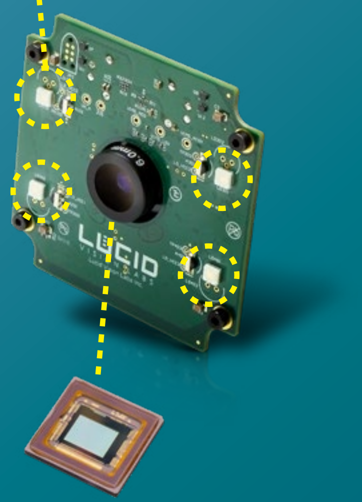
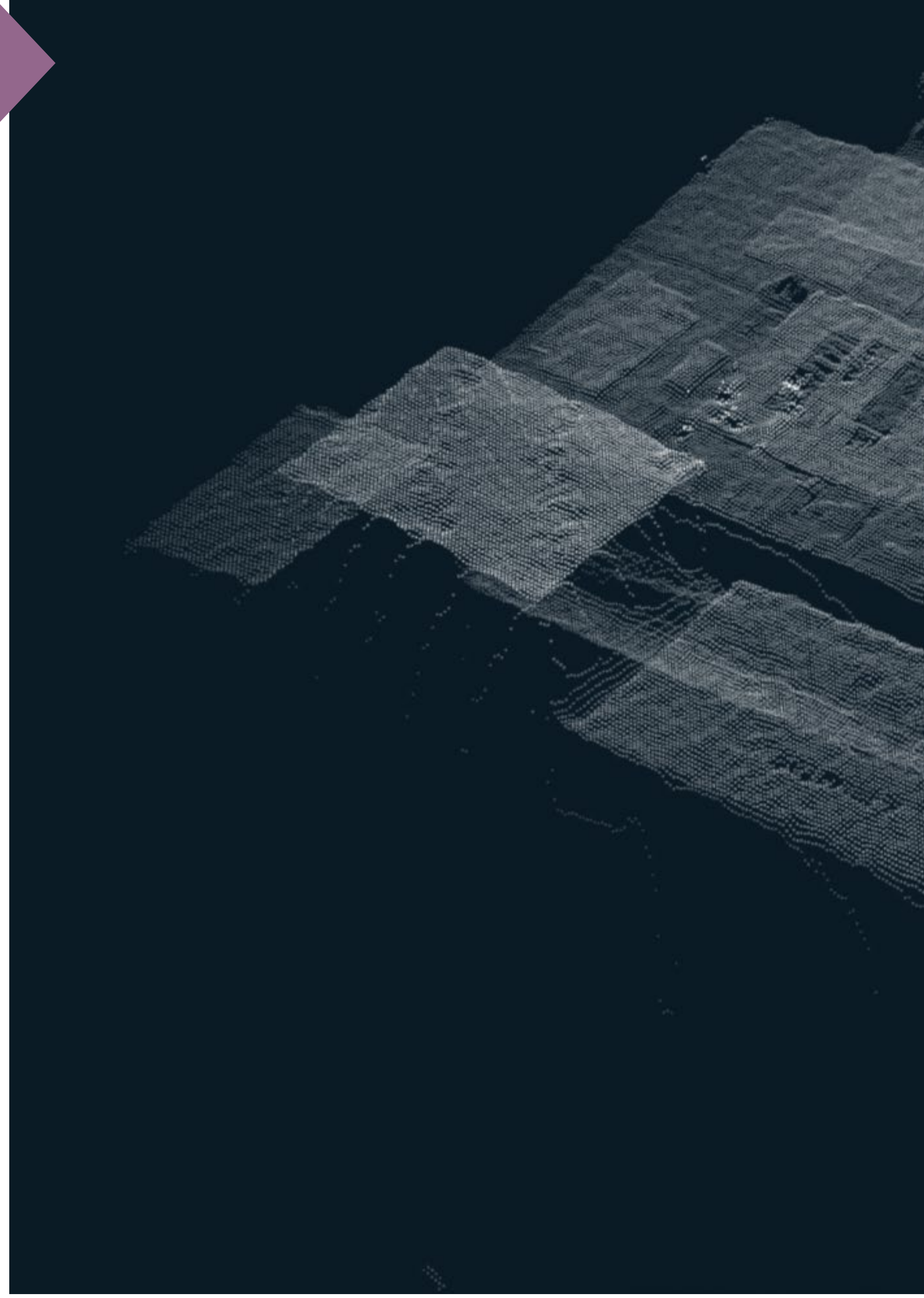


figure 12: Different application environments present different lighting challenges from multiple light sources. It is important to understand your application's lighting and object materials before choosing a 3D technology. In this example of a warehouse, various light sources will affect your 3D imaging in different ways.

figure 14: Original image of cardboard boxes on pallet.

figure 14: The Helios camera can output a point cloud to the host PC. In this example a point cloud is created of a pallet of cardboard boxes at a distance of 2m.



Lasting Benefits

Time of flight (ToF) technology such as Sony's DepthSense sensor used in LUCID's Helios camera ***(figure 13)*** presents a new set of opportunities for those developing 3D imaging systems. First, because the system only uses a single camera, no calibration is required by the developer. Secondly, the system is much less affected by adverse lighting conditions compared to traditional passive stereo. Thirdly, the camera outputs point cloud data ***(figure 14)*** directly, offloading processing from the host PC. Lastly, the system is relatively low-cost being less expensive than high-performance active laser systems and comparable with projected laser light stereo systems.

With a range of 6m and an accuracy of approximately 5mm, LUCID's Helios camera will find many different applications. Such cameras may be used in systems that previously used stereo cameras to capture objects. 3D object detection applications include robotic pick and place machines where the Helios camera would replace a dual-head stereo vision system, lowering the size and weight of the system. The Helios ToF camera would also reduce the processing required on the PC to recognize and localize objects since point cloud data is already calculated in the Helios camera. For applications such as materials inspection that previously used pattern projectors to illuminate areas with little or no surface features, the **Helios camera could be used because its ToF technology only requires objects reflect NIR light; it does not require texture for correspondence matching.**

3D imaging technology will continue to play a powerful role in industrial applications. Depending on the application requirements, such as distance to target, level of accuracy and precision, environmental lighting, and overall costs, different 3D technologies will perform better than others. For applications that demand real-time object detection, ToF technology enhances 3D depth performance and simplifies application complexity over traditional stereo vision camera systems. The Helios ToF camera featuring Sony's DepthSense sensor further improves the accuracy, precision, and speed of ToF technology, opening up new 3D possibilities for application designers and creating lasting benefits for the end customer.

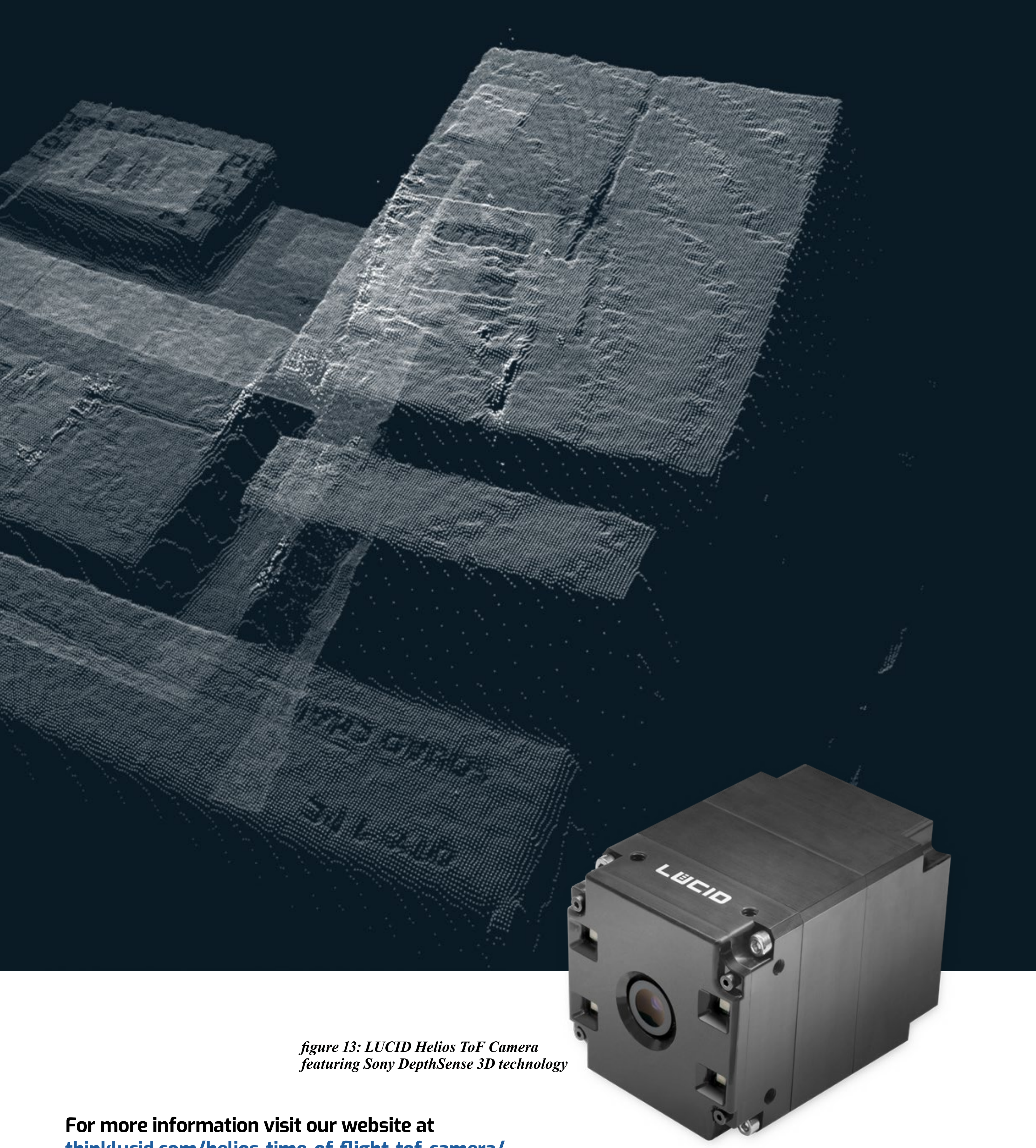


figure 13: LUCID Helios ToF Camera featuring Sony DepthSense 3D technology

For more information visit our website at
thinklucid.com/helios-time-of-flight-tof-camera/

LUCID
VISION LABS

LUCID Headquarters
LUCID Vision Labs, Inc.
130-13200 Delf Place, Richmond B.C.
Canada, V6V 2A2
sales@thinklucid.com
thinklucid.com
1-833-465-8243

Europe, Middle East, Africa
LUCID Vision Labs GmbH
Renntalstraße 14, 74360 Ilsfeld
Germany
sales.emea@thinklucid.com
+49 (0) 7062 97676 12

Asia Pacific
LUCID Vision Labs G.K
Eishin Bldg. 4F 3-6-1, Kanda-Ogawamachi,
Chiyoda-ku, Tokyo 101-0052
Japan
sales.apac@thinklucid.com
+81 3 5577 7915

Greater China
LUCID Vision Labs, China
sales.gc@thinklucid.com
thinklucid.cn

LUCID VISION LABS