THE USE OF HIGH PERFORMANCE STANDARD CMOS SENSORS FOR 3D VISION, DETECTION AND MEASUREMENT

By Ha Lan Do Thu, Marketing Manager for 3D Imaging at Teledyne e2v

Overview

3D imaging technology appeared several decades ago but the first products were only commercialized in the 2000s, when many big studio movies were released in 3D, using the latest HD video cameras. Since then, the field has evolved in leaps and bounds in terms of speed, accuracy and 3D imaging resolution and it has been extensively adopted not only in consumer markets but also in the machine vision industry.

With the Industry 4.0 revolution upon us, the need for 3D vision is increasing due to the limitations of 2D vision in terms of accuracy and distance measurement for complex object recognition, dimensioning and also limitations in complex interaction situations such as human/robot cohabitation.

3D vision increases the autonomy and effectiveness of robots/machine systems in the factory automation market as it's essential for higher accuracy quality inspection, reverse engineering and object dimensioning where 2D vision is limited. In addition, the use of visions system guided robotics is growing and requires 3D vision for better remote guidance, obstacle recognition and accurate moving.

3D vision also protects workers in factories from intensive human/machine interactions, with systems preventing and resolving dangerous situations, and with surveillance systems that are able to count and differentiate people from robots or objects.

3D vision is influencing society with its ability to make safer, better performing and more effective assistance systems for end-users. For instance, 3D vision is an enabler for advanced automotive assistance driver systems in autonomous vehicles, collaborative robots etc.

When considering scanning barcodes and OCR, 2D imaging is nevertheless here to stay. It plays an integral role in factories and warehouses, where its deployment is rising through block-chain adoption, as well as the e-commerce boom, which is driving spectacular growth in fulfilment centers and transportation. Teledyne e2v has unique 2D imaging products dedicated to code scanning such as the Topaz sensor family with performance and features to achieve high scanning rates and reliability.

Several techniques and technologies exist in order to get a 3D image. The main ones are:

- **Stereo vision** - two cameras are used, mounted with different perspectives of an object, and calibration techniques are used to align pixel information between the cameras and extract depth information. This is similar to how our brains work to visually measure distance.
- **Structured light** - a known light pattern is projected onto an object and depth information is calculated by the way the pattern is distorted around the object.

- **Laser triangulation** - laser triangulation systems enable 3-dimensional measurements to be captured by pairing a camera with a laser source. Due to a known angular offset between the laser and the camera, the system measures the geometrical offset of a laser line (whose value is related to the height of the object) using trigonometry. This technique is based on the scanning of the object
- **Time-of-Flight** - A light source is synchronized with an image sensor to calculate distance based on the time between the pulse of light and the reflected light back onto the sensor.

Each technique has its pros and cons and so, depending on the targeted applications (especially the distance range and the depth accuracy requirements), some are more suitable than others. A relative comparison is done in **Table 1**.

Performances	Stereo Vision	Structured Light	Laser Triangulation	Time of Flight (ToF)
Range	Limited (2 m to 5 m) Could be improved by combining with a light source	Scalable (cm to 3-5 m)	Short & Limited (cms to 1 m)	Scalable (<50 cm to 20-50 m)
Depth Accuracy	Low (cm) Could be improved by combining with a light source	High (µm to cm)	Very high (µm)	Medium (mm to cm)
Response Time	Medium	Slow	Slow	Fast
Software Complexity	High	Medium	High	Low
Low-Light Performance	Weak	Light source dependent (IR or visible)	Light source dependent (IR or visible)	Good (IR, laser, LED)
Bright-Light Performance	Good	Medium	Medium	Good
Changing Light Conditions	Weak	Medium	Medium	Good
Compactness	Low	Low	Medium	High
Manufacturing & Calibration efforts	High	High	High	Low
Material Cost	Low	High/Medium	High/Medium	Medium
Field-of-View	Limited (fixed and calibrated)	Limited (fixed and calibrated)	Limited (fixed and calibrated)	Scalable (large FoV possible, source, sensor resolution, distance dependent)

Table 1. ***3D Imaging techniques 'top-level' comparison***

Although it still represents a small part of the vision systems used in factory automation and warehouses, currently more and more 3D systems are deployed based on 3D stereo vision, structured light cameras or laser displacement. These systems operate at fixed working distances and require significant calibration efforts for specific areas of detection. A Time-of-Flight system overcomes these challenges and gives more flexibility from an application point of view but most of them are still limited in image resolution.

Teledyne e2v has a track record of successes in machine vision, both in line scan cameras and area scan sensors, and has now built a unique platform dedicated to 3D imaging. This will support the latest industrial applications such as vision guided robotics, logistics automated guided vehicles (AGV), factory surveillance and safety, handheld scanners as well as outdoor applications. Teledyne e2v aims to provide a consistent offer based on several 3D techniques that will meet the applications requirements of the customer.

Beyond 2D, 3D vision for high speed and accurate inspection to increase productivity

Factories are automated to increase productivity and gain time and money in all the cycles of product inspection and inventory. To optimize these factors, machines with vision systems need to operate at higher speeds and with better performance.

Subsequently, whereas 2D vision is limited, 3D vision is extensively deployed for higher accuracy quality inspection, reverse engineering or object dimensioning. The laser triangulation technique is commonly used in in these types of applications because of the high resolution required on the three axis, therefore demanding very high speed sensors.

3D Inspection | **Measuring Shapes** | **Volume Inspection (Logistics)**

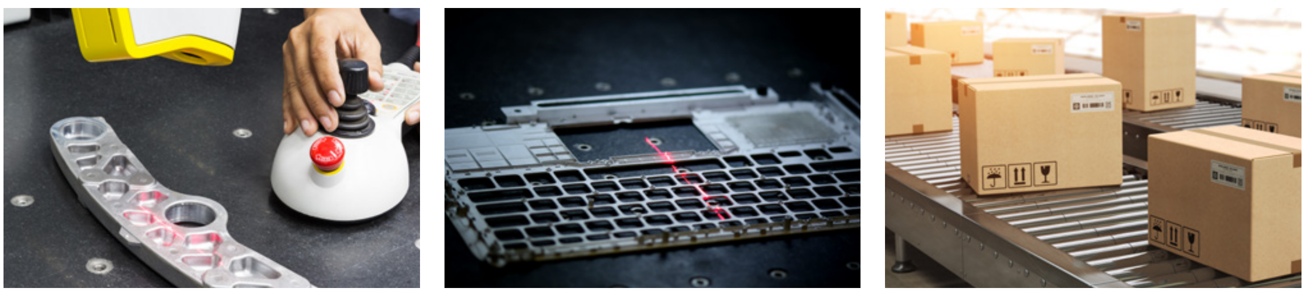

Figure 1. ***Laser triangulation applications examples***

For more than 10 years, Teledyne e2v has been working with market leaders in 3D laser triangulation and developing custom sensors. In 2019, Teledyne e2v launched Flash, a family of off-the-shelf sensors to meet the challenging demands of high speed volume measurement and inspection to comply with high speed production lines.

The Flash CMOS sensors outstandingly combine a resolution of 4,096 x 1,024 pixels and 2,048 x 1,024 pixels with a respective frame rate of 1,800 fps and 1,500 fps (8 bits) and a respective readout speed of 61.4 Gbps and 25.6 Gbps in a standard optical format (APS-like optics and C-Mount respectively).

Designed to enable easy and cost-effective integration for camera makers, the new sensors include a wide range of application-based features enabling high flexibility in use such as:

- High Dynamic Range modes with up to 100 dB that enables the measurement and inspection of both highly-reflective surfaces and dark areas in the same image through a HDR mode up to 100 dB
- Multiple Region of Interest mode that allows a perfect trade-off between profiling rate and range/resolution in the height measurement
- Frame to frame 'hot' changes mode of some parameters that enables flexibility and real-time adaptation to the environment conditions
- Different trigger modes that allow a perfect adaptation to the speed of production lines

The main benefits for the industrial market are:
- **A higher production rate** - due to its very high frame rate combined with the 2k or 4k horizontal resolution
- **Cost effective system** - the Flash family provides a high horizontal resolution with just the required vertical resolution (compared with standard form factors) to reduce silicon size and so lower sensor and system costs
- **High flexibility in use** - such as allowing a real-time adaptation to the environmental conditions and to the speed of the production lines thanks to its features like the single capture HDR.

Beyond 2D, 3D vision to ease autonomy and effectiveness of factories

In order to improve the autonomy and effectiveness of factory automation, the use of autonomous guided robots is growing. As factories and warehouses become increasingly automated, working stations are intensively collaborative between humans and machines making prevention and safety even more essential.

In such unpredictable working environments, Time-of-Flight (ToF) systems are ideal due to their real-time 3D information and decision capability and the ability to detect objects or people within fast scenes. There are two methods when talking about the Time-of-Flight technique: direct Time-of-Flight and indirect Time-of-Flight:

- **Direct ToF** - the system measures the distance by directly calculating the time it takes for the light to bounce off an object
- **Indirect ToF** - the system measures the distance by calculating the phase differences of the light waves when they are emitted to the object and when they bounce back. This allows the system to build a 3D map.

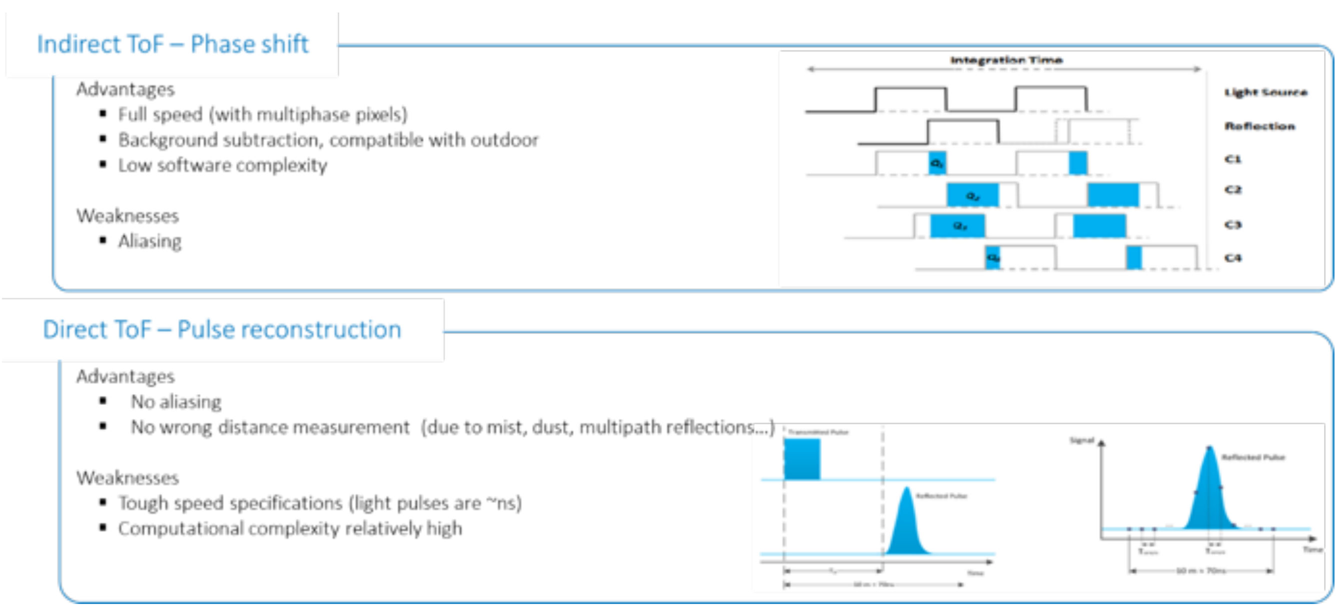

Figure 2. ***Pros and Cons of Direct ToF and Indirect ToF***

Teledyne e2v offers solutions for both direct Time-of-Flight systems through custom sensors and indirect Time-of-Flight systems through a complete portfolio of standard products.

Teledyne e2v has been working for over 10 years on Time-of-Flight technologies and solutions and has strong capabilities with an innovative Time-of-Flight pixel, which is based on a track record of custom sensors developed for leaders in automation robotics automotive, and surveillance.

We designed our standard ToF sensors with the following key differentiators that our customers are seeking:
- High spatial resolution sensors enabling a large field of view coverage and high angular resolution
- Fast and real-time 3D detection without any motion artefact and over 30 fps depth map at full resolution
- Both short and mid, long distance detection and management
- Ultra-reliable distance measurement in any condition: excellent precision and robust to ambient light conditions and multiple systems operation

Time-of-Flight to handle complex environment operations

Compared to others 3D techniques, the ToF system is fast, simple, inexpensive and has excellent 3D performances at mid and long range.

As previously mentioned, a ToF system is highly flexible, so a system could be mounted in a moving robot as it doesn't need to be calibrated depending on the setup (e.g. factory calibration). The system can also be adapted to any operational setup making a ToF system well suited to being operated in complex environments and conditions.

Below are some examples of factory and warehouse applications where Time-of-Flight is the best candidate.

Figure 3. ***ToF application examples***

- Measuring the size and volume of a box or package - ToF imaging systems can measure the size and volume of a box/package in order to get the most efficient palletization or conditioning (packing and truckload optimization) or for efficient product lines
- Goods identification for intelligent and efficient warehouse management - ToF imaging systems detect the goods or dimensions of a package in a shorter time than 2D conventional image processing
- Pick and place - ToF imaging systems detect and identify the right object to be picked and placed, in the right position, with a very high level of accuracy and in shorter times than conventional 2D image processing.

Helping robots to navigate autonomously and safely in factory environment in another application where ToF systems are beneficial compared to other 3D techniques.

Figure 4. ***Robot navigation application example***

Today, most navigation systems use a traditional Lidar scanner to detect objects but there are some cons and more and more solutions are emerging with alternatives technologies such as indirect ToF. Compared to the traditional Lidar scanner, indirect ToF systems can provide real time image information with fast response times and no motion artefacts. They are also more robust and compact as no mechanical part is involved and cheaper since it's a solid-state design using less power and less computation.

Parameters	Traditional Lidar	Indirect Time of Flight (ToF)
Range	Long (Up to 70 m-100 m)	Scalable (Short distance: 30-50 cm to 5 m-10 m; Long distance: >10 m-100 m)
Precision	Medium: Laser scanning, mm	Medium: mm, cm
Resolution	Low	Medium-High: Up to 1.3 MP: 1280x1024 (state-of-the-art)
Speed	Slow: 20 fps	High: >60 fps for 3D depth map
2D and 3D images	No: No image, coordinates reconstruction by computation	Yes: 2D greyscale mage (CMOS sensors) + 3D Map
Motion Artifacts Removal	No: One single point scanning of the object, may cause artifacts or miss the object	Yes: capture of moving objects instantaneously (in one frame)
Field-of-View	Large in horizontal (laser scanning up to 360°) Low in vertical	Medium: Can be large (> 120°) depending on sensor resolution, precision, illumination trade-off
Response Time: 3D map	Low: Scanning before 3D reconstruction	High: Real time capture of the scene at whole field of view
Compactness of the System	Medium	High: Solid state system
Robustness of the System	Medium: Mechanical parts	High: solid state system
Robustness to the outdoor environment (changing-light, sunlight...)	High	Medium
Price	Expensive	Cheaper

Table 2. ***Traditional Lidar vs Indirect ToF 'top level' comparison***

In the aforementioned applications, a ToF system tackles many challenges. The system needs to handle both short and mid, long range distances starting from 10 meters as well as needing to be fast without any motion artefacts (today, most ToF solutions on the market are focused on short ranges up to 5-6 meters). In addition, the system could be exposed to several changing lighting conditions as the robot is moving throughout the factory/warehouse and needs to operate without any interference from others robots in the same area.

To handle all of these challenging conditions, Teledyne e2v introduced the Hydra3D sensor in July 2020. This best in class ToF sensor, has an innovative pixel and highly flexible configuration delivering very high dynamic range, which is quite intrinsic to ToF applications (a combination of a large range of both objects reflectivity and distance). Hydra3D has a 832 x 600 pixel resolution incorporating a 10µm, three tap cutting-edge pixel, which allows the highest levels of 3D performance, including high depth resolution, high speed and flexible operation conditions.

SENSOR FEATURES	CUSTOMER BENEFITS
Spatial resolution of 832 x 600 pixels in both 2D and 3D LGA Ceramic package 24 x 22 mm Compatible with 2/3" optics	**Large Field-of-View with good angular resolution in a compact sensor**
Excellent temporal precision due to fast transfer time as low as 20 ns Three-tap, cutting-edge pixel Multi-system management feature embedded on-chip More than 30 fps depth map	**Real-time decision making combined with reliable 3D detection, without motion artefacts and interference with other systems**
Short and long-distance range handling (>10 m) High flexibility to trade-off distance range, object reflectivity, frame rate, etc. combined with powerful non-destructive readout HDR Robust to ambient light and challenging environments	**Outstanding adaptability to all environments, including outdoors, with very high dynamic range management**

Hydra3D also has an innovative multi-tap pixel that is able to simultaneously acquire the three phases required to build the 3D image, enabling accurate 3D information to be captured of fast moving scenes without any motion artefacts.

The three-tap pixel provides the ability to capture and store the information of the different phases needed to reconstruct the 3D map in one single frame, instead of the multiple frames required with other indirect ToF technologies. This is particularly powerful in avoiding any motion artefacts in fast moving scenes. It's similar to comparing a global shutter mode to a rolling shutter mode in 2D vision. The three-tap pixel can capture all phases with a single train of light pulses, making optimal use of the light power.

In the example described, three images are necessary (two phases + the background). If using a 1-tap pixel sensor (which is the most common today), it is required to expose and readout the multiple phases sequentially. Shooting the light once to get phase 0, another time to get phase 1 and then run a third acquisition without light and readout to acquire the background. Only then computing the 3D image. Therefore, if there is a moving object, motion artefacts will appear since the object will be in a different position within each capture. Also, the light needs to be shot twice, once for each phase.

Using a multiple-tap pixel sensor, for instance a 3-tap pixel in the case above, all exposures and readouts are made in an interleaved way, so that all phases are acquired virtually in parallel, minimizing motion artefacts. Furthermore, since the phases can be captured with a single train of light pulses, it reduces the average light power, which is important both from an eye-safety and power consumption point of view.

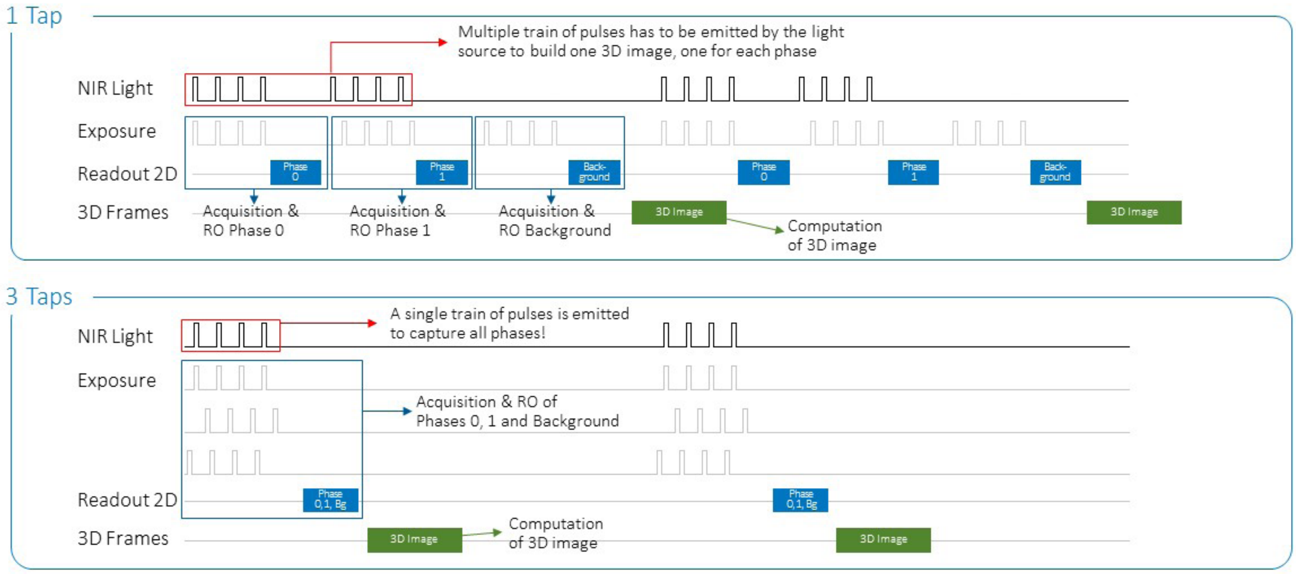

Figure 5. ***One-tap pixel vs Three-tap pixel illustration***

Note that the difference between a motion artefact and motion blur is similar to the difference between rolling and global shutters in 2D vision. Motion blur can slightly warp the objects which are moving fast, but doesn't provide false information, while motion artefacts can drastically change the appearance of the objects and provide false measurement, which can have big consequences in some applications.

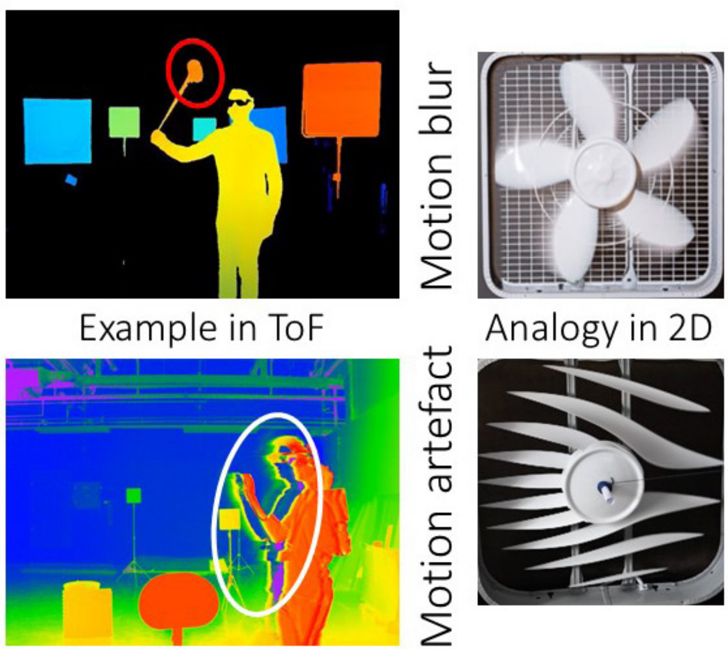

Figure 6. ***Motion blur vs Motion artefact***

Other benefits of the Hydra3D sensor are its flexibility in configuration and its multiple on-chip features. For instance, a powerful on-chip HDR feature combined with a high frame rate, and a flexible configuration allows you to trade-off between distance range, object reflectivity, frame rate, etc. as well as being robust to the ambient light. In a single trigger, Hydra3D initiates a sequence of acquisitions and readouts, easily programmable, resulting in a very powerful tool to adapt to the conditions of each application.

Multiple measuring sequences allow measurement at different distance ranges, or with different precision, or to perform 2D captures, all within a single trigger. And on top of that, the sequence is changeable from frame-to-frame, live, without halting the sensor.

The unique on-chip feature for robust multiple ToF system operation enables unsynchronized systems to work simultaneously without interfering with each other. Since ToF requires active illumination, one system can suffer from interferences caused by the light emitted by another system working simultaneously in the same area that can lead to incorrect distance measurements.

Key characteristics of the Hydra3D sensor:

	SENSOR CHARACTERISTICS			
Resolution – pixels	832 x 600			
Aspect Ratio	4 : 3			
Size Type	2/3" (10.3 mm diagonal)			
Pixel Type / size – square	Three-tap global shutter – Gated global shutter / 10 µm			
Maximum frame rate @ 12 bits	416.7 fps[1]			
FFxQE – %, @ 850 / 940 nm	37 / 19[2]			
Transfer time – ns	Down to 20			
Readout noise – e- RMS	2.5			
Linearity: LEmin / LEmax – %	-1 / +1			
	Node A	Node B	Node C	A+B+C
Full well capacity – e-	10,000	10,000	10,000	30,000
Temporal noise – e-	10	10	10	17.3
Dynamic Range[3] – dB	60	60	60	64.7

1. Considering only readout. Exposure is not concurrent
2. In 2D greyscale mode
3. Single readout, 2D greyscale mode

Time-of-Flight in the age of CCD adoption

Time-of-Flight image processing is much more complex than a conventional 2D vision system. It involves optics and a light system that will depend on several parameters (such as the sensor or the field-of-view, in-factory calibration, specific sensor configuration) to fit perfectly with the application requirements.

We often compare Time-of-Flight sensors to CCD sensors since the complexity of the set-up requires a significant effort at system level to integrate the sensor and so the need to master the application set-up becomes critical.

Based on our expertise in ToF CMOS sensors and from working closely with our customers, Teledyne e2v has gained solid experience in assessing the challenges inherent to ToF systems. In order to help our customers to shorten their time-to-market and get the best ToF system to fit their application requirements, we provide technology solutions ranging from CMOS image sensors and customized camera modules, right up to full system integration support. This includes a reference design of our evaluation platform, light & optics assessment, eye safety consideration, modelisation and simulations, algorithms and calibrations.

Figure 7. ***Teledyne-e2v ToF expertise and capabilities up to system level***

Summary

As factories and distribution warehouse are increasingly automated, there is an increasing need for effective and autonomous industrial systems, particularly 3D vision for guided robots and machines (for object recognition, navigation, high speed and accuracy). Several 3D techniques exist, each with pros and cons, with the preferred technology, very much depending on the requirements of the application. All of these techniques are demanding and require high performance sensors with complex features.

Teledyne e2v has a successful track record with market leaders and offers a broad range of unique solutions, including 3D vision, to serve industrial markets such as factory automation, logistics and metrology applications. Our unique expertise in high performance CMOS image sensors (incorporating cutting-edge pixels and special features) coupled with over 10 years solid experience in ToF systems, enables us to help our customers to overcome the current challenges of 3D vision.