

Time-of-Flight Camera – An Introduction

Larry Li

Sensing Solutions

1. Introduction

3D Time-of-Flight (TOF) technology is revolutionizing the machine vision industry by providing 3D imaging using a low-cost CMOS pixel array together with an active modulated light source. Compact construction, easy-of-use, together with high accuracy and frame-rate makes TOF cameras an attractive solution for a wide range of applications. In this article, we will cover the basics of TOF operation, and compare TOF with other 2D/3D vision technologies. Then various applications that benefit from TOF sensing, such as gesturing and 3D scanning and printing, are explored. Finally, resources that help readers get started with Texas Instruments' 3D TOF solution are provided.

2. Theory of Operation

A 3D time-of-flight (TOF) camera works by illuminating the scene with a modulated light source, and observing the reflected light. The phase shift between the illumination and the reflection is measured and translated to distance. Figure 1 illustrates the basic TOF concept. Typically, the illumination is from a solid-state laser or a LED operating in the near-infrared range (~850nm) invisible to the human eyes. An imaging sensor designed to respond to the same spectrum receives the light and converts the photonic energy to electrical current. Note that the light entering the sensor has an *ambient* component and a *reflected* component. Distance (depth) information is only embedded in the reflected component. Therefore, high ambient component reduces the signal to noise ratio (SNR).

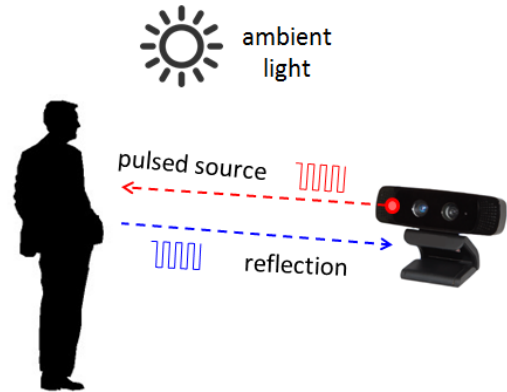


Figure 1: 3D time-of-flight camera operation.

To detect phase shifts between the illumination and the reflection, the light source is **pulsed** or modulated by a **continuous-wave (CW)**, source, typically a sinusoid or square wave. Square wave modulation is more common because it can be easily realized using digital circuits [5].

Pulsed modulation can be achieved by integrating photoelectrons from the reflected light, or by starting a fast counter at the first detection of the reflection. The latter requires a fast photo-detector, usually a *single-photon avalanche diode* (SPAD). This counting approach necessitates fast electronics, since achieving 1 millimeter accuracy requires timing a pulse of 6.6 picoseconds in duration. This level of accuracy is nearly impossible to achieve in silicon at room temperature [1].

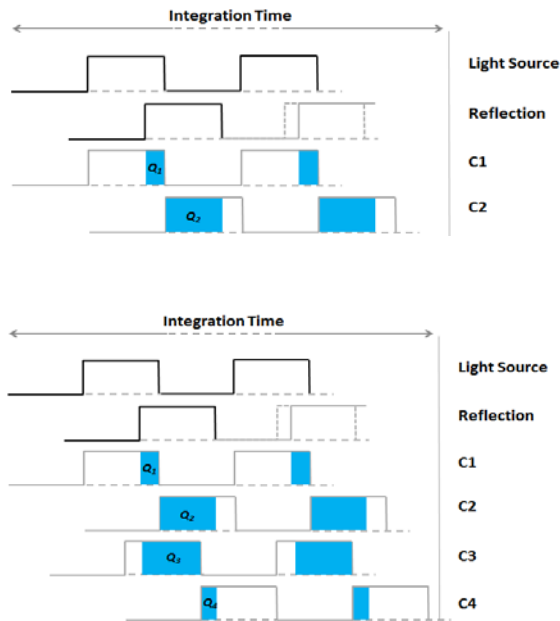


Figure 2: Two time-of-flight methods: pulsed (top) and continuous-wave (bottom).

The *pulsed* method is straightforward. The light source illuminates for a brief period (Δt), and the reflected energy is sampled at every pixel, in parallel, using two out-of-phase windows, C_1 and C_2 , with the same Δt . Electrical charges accumulated during these samples, Q_1 and Q_2 , are measured and used to compute distance using the formula:

$$d = \frac{1}{2} c \Delta t \left(\frac{Q_2}{Q_1 + Q_2} \right). \quad \text{Eq. 1}$$

In contrast, the CW method takes multiple samples per measurement, with each sample phase-stepped by 90 degrees, for a total of four samples. Using this technique, the phase angle between illumination and reflection, φ , and the distance, d , can be calculated by

$$\varphi = \arctan \left(\frac{Q_3 - Q_4}{Q_1 - Q_2} \right), \quad \text{Eq. 2}$$

$$d = \frac{c}{4\pi f} \varphi. \quad \text{Eq. 3}$$

It follows that the measured pixel intensity (A) and offset (B) can be computed by:

$$A = \frac{\sqrt{(Q_1 - Q_2)^2 + (Q_3 - Q_4)^2}}{2}, \quad \text{Eq. 4}$$

$$B = \frac{Q_1 + Q_2 + Q_3 + Q_4}{4}. \quad \text{Eq. 5}$$

In all of the equations, c is the speed-of-light constant.

At first glance, the complexity of the CW method, as compared to the pulsed method, may seem unjustified, but a closer look at the CW equations reveals that the terms, $(Q_3 - Q_4)$ and $(Q_1 - Q_2)$ reduces the effect of constant offset from the measurements. Furthermore, the quotient in the phase equation reduces the effects of constant gains from the distance measurements, such as system amplification and attenuation, or the reflected intensity. These are desirable properties.

The reflected amplitude (A) and offset (B) do have an impact the depth measurement accuracy. The depth measurement variance can be approximated by:

$$\sigma = \frac{c}{4\sqrt{2}\pi f} \cdot \frac{\sqrt{A+B}}{c_d A} \quad \text{Eq. 6}$$

The modulation contrast, c_d , describes how well the TOF sensor separates and collects the photoelectrons. The reflected amplitude, A , is a function of the optical power. The offset, B , is a function of the ambient light and residual system offset. One may infer from Equation 6 that high amplitude, high modulation frequency and high modulation contrast will increase accuracy; while high offset can lead to saturation and reduce accuracy.

At high frequency, the modulation contrast can begin to attenuate due to the physical property of the silicon. This puts a practical upper limit on the modulation frequency. TOF sensors with high roll-off frequency generally can deliver higher accuracy.

The fact that the CW measurement is based on phase, which wraps around every 2π , means the distance will also have an aliasing distance. The distance where aliasing occurs is called the ambiguity distance, d_{amb} , and is defined as:

$$d_{amb} = \frac{c}{2f} \quad \text{Eq. 7}$$

Since the distance wraps, d_{amb} is also the maximum measurable distance. If one wishes to extend the measurable distance, one may reduce the modulation frequency, but at the cost of reduced accuracy, as according to Equation 6.

Instead of accepting this compromise, advanced TOF systems deploy multi-frequency techniques to extend the distance without reducing the modulation frequency. Multi-frequency techniques work by adding one or more modulation frequencies to the mix. Each modulation frequency will have a different ambiguity distance, but true location is the one where the different frequencies agree. The frequency of when the two modulations agree, called the *beat frequency*, is usually lower, and corresponds to a much longer ambiguity distance. The dual-frequency concept is illustrated below.

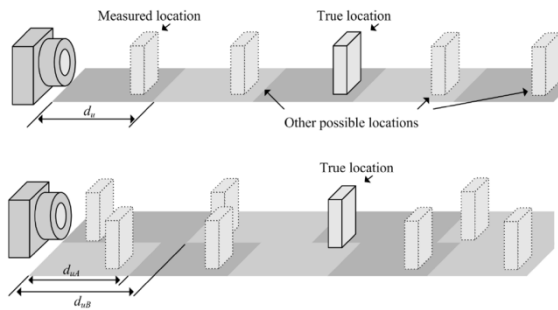


Figure 3: Extending distance using a multi-frequency technique [6].

3. Point Cloud

In TOF sensors, distance is measured for every pixel in a 2D addressable array, resulting in a **depth map**. A depth map is a collection of 3D points (each point also known as a **voxel**). As an example, a QVGA sensor will have a depth map of 320 x 240 voxels. 2D representation of a depth map is a gray-scale image, as is illustrated by the soda cans example in Figure 4—the brighter the intensity, the closer the voxel. Figure 4 shows the depth map of a group of soda cans.



Figure 4: Depth map of soda cans.

Alternatively, a depth map can be rendered in a three-dimensional space as a collection of points, or **point-cloud**. The 3D points can be mathematically connected to form a mesh onto which a texture surface can be mapped. If the texture is from a real-time color image of the same subject, a life-like 3D rendering of the subject will emerge, as is illustrated by the avatar in Figure 5. One may be able to rotate the avatar to view different perspectives.

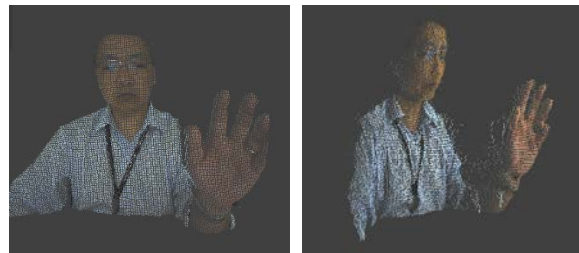


Figure 5: Avatar formed from point-cloud.

4. Other Vision Technologies

Time-of-flight technology is not the only vision technology available. In this section, we will compare TOF to the classical 2D machine vision and other 3D vision technologies. A table summarizing the comparison is included at the end of this section.

2D Machine Vision

Most machine vision systems deployed today are 2D, a cost-effective approach when lighting is closely controlled. They are well-suited for inspection applications where defects are detected using well-known image processing techniques, such as edge detection, template matching and morphology open/close. These algorithms extract critical feature parameters that are compared to a database for pass-fail determination. To detect defects along the z-axis, an additional 1D sensor or 3D vision is often deployed.

2D vision could be used in unstructured environment as well with the aid of advanced image processing algorithms to get around complications caused by varying illumination and shading conditions. Take the images in Figure 6 for example. These images are from the same face, but under very different lighting. The shading differences can make face recognition difficult even for humans.

In contrast, computer recognition using point cloud data from TOF sensors is largely unaffected by shading, since illumination is provided by the TOF sensor itself, and the depth measurement is extracted from phase measurement, not image intensity.

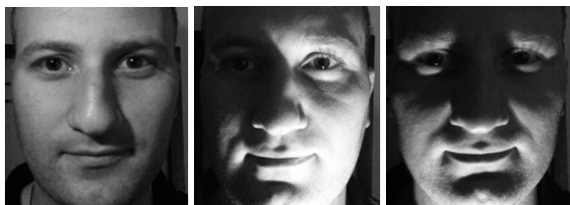
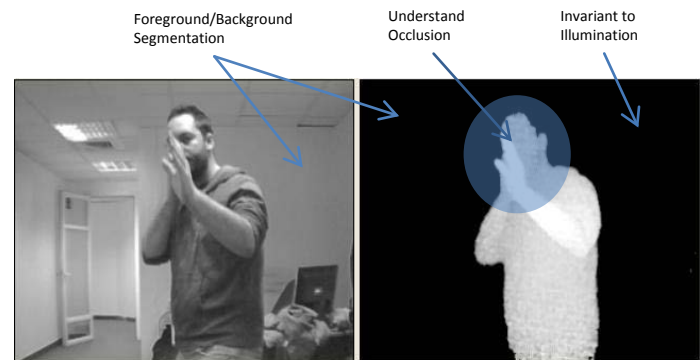


Figure 6: Same face, different shading.

3D Machine Vision

Robust 3D vision overcomes many problems of 2D vision, as the depth measurement can be used to easily separate foreground from background. This is particularly useful for scene understanding, where the first step is to segment the subject of interest (foreground) from other parts of the image (background).

Gesture recognition, for example, involves scene understanding. Using distance as a discriminator, a TOF sensor enables separation of the face, hands, and fingers from the rest of the image, so gesture recognition can be achieved with high confidence.



*Used with permission

Figure 7: Advantages of 3D vision over 2D.

In the next two subsections we will compare the TOF technology with two other 3D vision technologies: **stereo vision** and **structured-light**.

➤ **Stereo Vision vs. TOF**

Stereo vision generally uses two cameras separated by a distance, in a physical arrangement similar to the human eyes. Given a point-like object in space, the camera separation will lead to measurable disparity of the object positions in the two camera images. Using a simple pin-hole camera model, the object position in each image can be computed, which we will represent them by α and β . With these angles, the depth, z , can be computed.

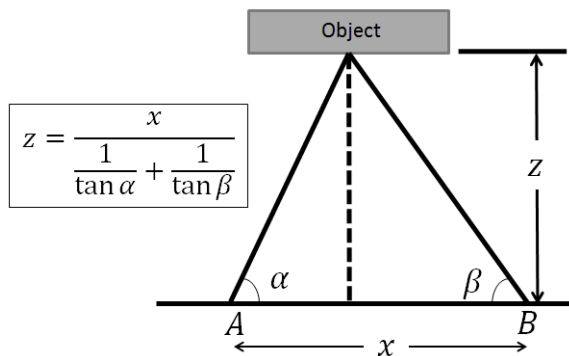


Figure 8: Stereopsis—depth through disparity measurement.

A major challenge in stereo vision is solving the **correspondence problem**: *giving a point in one image, how to find the same point in the other camera?* Until the correspondence can be established, disparity, and therefore depth, cannot be accurately determined. Solving the correspondence problem involves complex, computationally intensive algorithms for feature extraction and matching. Feature extraction and matching also require sufficient intensity and color variation in the image for robust correlation. This requirement renders stereo vision less effective if the subject lacks these variations—for example, measuring the distance to a uniformly colored wall. TOF sensing does not have this limitation because it

does not depend on color or texture to measure the distance.

In stereo vision, **the depth resolution error is a quadratic function of the distance**. By comparison, a TOF sensor, which works off reflected light, is also sensitive to distance. However, the difference is that, for TOF this shortcoming is remedied by increasing the illumination energy when necessary; and the intensity information is used by TOF as a “confidence” metric to maximize accuracy using Kalman filter-like techniques.

Stereo vision has some advantages. The **implementation cost is very low**, as most common off-the-shelf cameras can be used. Also, the human-like physical configuration makes stereo vision **well-suited for capturing images for intuitive presentation to humans**, so that both humans and machines are looking at the same images.

➤ **Structured-Light vs. TOF**

Structured-Light works by projecting known patterns onto the subject and inspecting the pattern distortion [4]. Successive projections of coded or phase-shifted patterns are often required to extract a single depth frame, which leads to a **lower frame rate**. Low frame rate means the **subject must remain relatively still** during the projection sequence to avoid blurring. The reflected pattern is sensitive to optical interference from the environment; therefore, structured-light tends to be better suited for indoor applications. A major advantage of structured-light is that it can achieve relatively high spatial (X-Y) resolution by using off-the-shelf DLP projectors and HD color cameras. Figure 9 shows the structured-light concept.

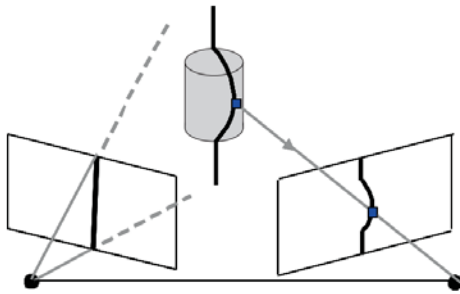


Figure 9: Structured-light concept.

By comparison, TOF is less sensitive to mechanical alignment and environmental lighting conditions,

and is more mechanically compact. The current TOF technology has lower resolution than today's structured-light, but is rapidly improving.

The comparison of TOF camera with stereo vision and structured-light is summarized in Table 1. The key takeaway is that ***TOF is a cost-effective, mechanically compact depth imaging solution unaffected by varying environmental illumination and vastly simplifies the figure-ground separation commonly required in scene understanding.*** This powerful combination makes TOF sensor well-suited for a wide variety of applications.

Table 1: Comparison of 3D Imaging Technologies

CONSIDERATIONS	STEREO VISION	STRUCTURED-LIGHT	TIME-OF-FLIGHT (TOF)
Software Complexity	High	Medium	Low
Material Cost	Low	High	Medium
Compactness	Low	High	Low
Response Time	Medium	Slow	Fast
Depth Accuracy	Low	High	Medium
Low-Light Performance	Weak	Good	Good
Bright-Light Performance	Good	Weak	Good
Power Consumption	Low	Medium	Scalable
Range	Limited	Scalable	Scalable
APPLICATIONS			
Game		X	X
3D Movies	X		
3D Scanning		X	X
User Interface Control			X
Augmented Reality	X		X

5. Applications

TOF technology can be applied to applications from automotive to industrial to healthcare, to smart advertising, gaming and entertainment. A TOF sensor could also serve as an excellent input device to both stationary and portable computing devices. In automotive, TOF sensors could enable autonomous driving and increased surrounding awareness for safety. In the industrial segment, TOF sensors could be used as HMI, and for enforcing

safety envelopes in automation cells where humans and robots may need to work in close proximity. In smart advertising, using TOF sensors as gesture input and human recognition, digital signage could become highly interactive, targeting media contents to the specific live audience. In healthcare, gesture recognition offers non-contact human-machine interactions, fostering more sanitary operating environment. The gesturing capability is particularly

well-suited for consumer electronics, particularly in gaming, portable computing, and home entertainment. TOF sensors natural interface provides an intuitive gaming interface for first-person video games. This same interface could also replace remote controls, mice and touch screens.

Generally speaking, TOF applications can be categorized into Gesture and Non-Gesture. Gesture applications emphasize human interactions and speed; while non-gesture applications emphasize measurement accuracy.



Figure 10: TOF technology applies to a wide range of applications.

Gesture Applications

Gesture applications translate human movements (faces, hands, fingers or whole-body) into symbolic directives to command gaming consoles, smart televisions, or portable computing devices. For examples, channel surfing can be done by waving of hands, and presentation can be scrolled by using finger flickering. These applications usually require fast response time, low- to medium-range, centimeter-level accuracy and power consumption.





Figure 11: Gesture recognition using a 3D-TOF sensor and SoftKinetic iisu® middleware.

Non-Gesture Applications

TOF sensors can be used in non-gesture applications as well. For instance, in **automotive**, a TOF camera can increase safety by alerting the driver when it detects people and objects in the vicinity of the car, and in computer assisted driving. In **robotics and automation**, TOF sensors can help detect product defects and enforce safety envelopes required for humans and robots to work in close proximity. With **3D printing** rapidly becoming popular and affordable, TOF cameras can be used to perform **3D scanning** to enable “3D copier” capability. In all of these applications, spatial accuracy is important.



* Used with permission

6. How to Get Started

Texas Instruments 3D-TOF solution is based on the CW method, but has added many enhancements for power reduction and increased accuracy.

Hardware

The current generation is a 3-chip solution, comprised a TOF imaging sensor (OPT81x0), the analog front-end (VSP5324) and a TOF controller (OPT91xx). The solution is available also as a camera development kit (CDK). Figure 12 shows the system block diagram.

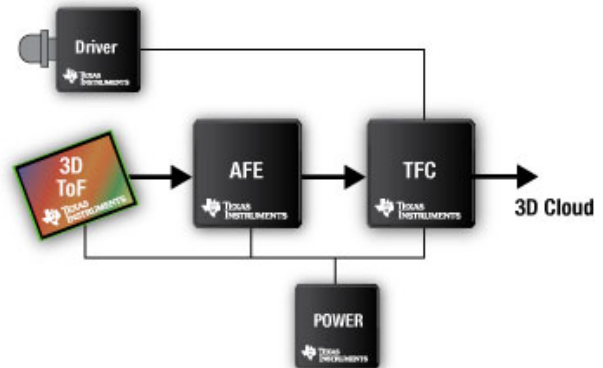


Figure 12: TI 3D-TOF chip set.

3D TOF Sensor Array

The 3D TOF sensor is an addressable CMOS pixel array based on the DepthSense™ technology [2][3], which supports high pixel modulation frequency (>50MHz), and up to 5x increase in signal-to-noise ratio (SNR). Through an advanced silicon fabrication process, the pixel device material is carefully tuned to respond to a specific optical spectrum (850nm-870nm) and to optimize the photoelectron collection and transport to the readout channels.

Analog Front-End (AFE)

The AFE (VSP5324) supports up to 4 differential inputs, each with a sample-and-hold front-end that also helps reject common-mode noise. A high-speed, low-power 12-bit ADC samples the input at up to 80 MSPS, and delivers the data serially over two differential serial data channels.

TOF Controller (TFC)

The TFC is a sophisticated state machine that synchronizes the operation of TOF sensor, AFE and illumination. It scans the pixels, calculates the depth for each pixel, and performs de-aliasing, de-noising, frequency tuning and temperature compensation. It also handles high-rate input and output, and serialization and deserialization.

Point cloud data from the TFC can be accessed by an external host through a variety of means. For webcam-like applications, the TFC can interface to a USB controller, such as the FX2, and enumerate as a USB Video Class (UVC) device. Alternatively, the TFC may interface with an embedded processor, such as OMAP™ and Sitara™ through a native camera interface (DVP/CAMIF) or a 4-lane SPI interface.

Different applications demand different illumination power, therefore, the power subsystem (PMIC) and laser/LED driver are intentionally kept as external components, so the solution is scalable. To complete the solution, Texas Instruments also offers a selection of PMIC and drivers. For more detail on Texas Instruments' TOF solution, please visit <http://www.ti.com/3dtof>.

References

- [1] E. Charbon, M. Fisburn, R. Walker, R. Henderson, C. Niclass, *SPAD-Based Sensors*.
- [2] D. Van Nieuwenhove, W. Van der Tempel, R. Grotjans, M. Kuijk, "Time-of-flight Optical Ranging Sensor based on a Current Assisted Photonic Demodulator", *Proceedings Symposium, IEEE/LEOS Benelux Chapter*, 2006, Eindhoven.
- [3] SoftKinetic DepthSense™ sensors, <http://www.softkinetic.com/en-us/products/depthsensesensors.aspx>.
- [4] J. Geng, "Structured-Light 3D Surface Imaging: a Tutorial", *Advances in Optics and Photonics* 3, 128-160 (2011).
- [5] M. Hansard, S. Lee, O. Choi, R. Horaud, "Time-of-Flight Cameras: Principles, Methods and Applications", *Springer Brief in Computer Science*, 2012.
- [6] A.P.P Jongenelen, D. G. Bailey, A.D. Payne, A.A. Dorrington, D.A. Carnegie, Analysis of Errors in ToF Range Imaging with Dual-Frequency Modulation, *IEEE Transactions on Instrumentation and Measurement*, Vo. 60, No. 5, May 2011, SSN 0018-9456.

Software

Texas Instruments has teamed up with several software vendors to provide state-of-the-art gesture recognition libraries or SDK that will work with the point cloud generated by the TI solution. There are also open-source libraries available with a rich set of features. The URLs for these software resources are listed below:

SoftKinetic: <http://www.softkinetic.com>

Point-Cloud Library: <http://www.pointclouds.org>

OpenCV: <http://www.opencv.org>

7. Summary

In this paper, we introduced the TOF technology and the theory of operation. We also compared TOF sensors with 2D machine vision and other 3D vision technologies and highlighted TOF sensors' differentiating advantages. We also explored a wide range of applications that TOF sensors enable or enhance. To help readers get started, we introduced TI 3D-TOF chipset and CDK, as well as third-party software resources.

For questions on the TI TOF products, please contact Ravi Ayyagari at rash@ti.com.

IMPORTANT NOTICE

Texas Instruments Incorporated and its subsidiaries (TI) reserve the right to make corrections, enhancements, improvements and other changes to its semiconductor products and services per JESD46, latest issue, and to discontinue any product or service per JESD48, latest issue. Buyers should obtain the latest relevant information before placing orders and should verify that such information is current and complete. All semiconductor products (also referred to herein as "components") are sold subject to TI's terms and conditions of sale supplied at the time of order acknowledgment.

TI warrants performance of its components to the specifications applicable at the time of sale, in accordance with the warranty in TI's terms and conditions of sale of semiconductor products. Testing and other quality control techniques are used to the extent TI deems necessary to support this warranty. Except where mandated by applicable law, testing of all parameters of each component is not necessarily performed.

TI assumes no liability for applications assistance or the design of Buyers' products. Buyers are responsible for their products and applications using TI components. To minimize the risks associated with Buyers' products and applications, Buyers should provide adequate design and operating safeguards.

TI does not warrant or represent that any license, either express or implied, is granted under any patent right, copyright, mask work right, or other intellectual property right relating to any combination, machine, or process in which TI components or services are used. Information published by TI regarding third-party products or services does not constitute a license to use such products or services or a warranty or endorsement thereof. Use of such information may require a license from a third party under the patents or other intellectual property of the third party, or a license from TI under the patents or other intellectual property of TI.

Reproduction of significant portions of TI information in TI data books or data sheets is permissible only if reproduction is without alteration and is accompanied by all associated warranties, conditions, limitations, and notices. TI is not responsible or liable for such altered documentation. Information of third parties may be subject to additional restrictions.

Resale of TI components or services with statements different from or beyond the parameters stated by TI for that component or service voids all express and any implied warranties for the associated TI component or service and is an unfair and deceptive business practice. TI is not responsible or liable for any such statements.

Buyer acknowledges and agrees that it is solely responsible for compliance with all legal, regulatory and safety-related requirements concerning its products, and any use of TI components in its applications, notwithstanding any applications-related information or support that may be provided by TI. Buyer represents and agrees that it has all the necessary expertise to create and implement safeguards which anticipate dangerous consequences of failures, monitor failures and their consequences, lessen the likelihood of failures that might cause harm and take appropriate remedial actions. Buyer will fully indemnify TI and its representatives against any damages arising out of the use of any TI components in safety-critical applications.

In some cases, TI components may be promoted specifically to facilitate safety-related applications. With such components, TI's goal is to help enable customers to design and create their own end-product solutions that meet applicable functional safety standards and requirements. Nonetheless, such components are subject to these terms.

No TI components are authorized for use in FDA Class III (or similar life-critical medical equipment) unless authorized officers of the parties have executed a special agreement specifically governing such use.

Only those TI components which TI has specifically designated as military grade or "enhanced plastic" are designed and intended for use in military/aerospace applications or environments. Buyer acknowledges and agrees that any military or aerospace use of TI components which have **not** been so designated is solely at the Buyer's risk, and that Buyer is solely responsible for compliance with all legal and regulatory requirements in connection with such use.

TI has specifically designated certain components as meeting ISO/TS16949 requirements, mainly for automotive use. In any case of use of non-designated products, TI will not be responsible for any failure to meet ISO/TS16949.

Products

Audio	www.ti.com/audio
Amplifiers	amplifier.ti.com
Data Converters	dataconverter.ti.com
DLP® Products	www.dlp.com
DSP	dsp.ti.com
Clocks and Timers	www.ti.com/clocks
Interface	interface.ti.com
Logic	logic.ti.com
Power Mgmt	power.ti.com
Microcontrollers	microcontroller.ti.com
RFID	www.ti-rfid.com
OMAP Applications Processors	www.ti.com/omap
Wireless Connectivity	www.ti.com/wirelessconnectivity

Applications

Automotive and Transportation	www.ti.com/automotive
Communications and Telecom	www.ti.com/communications
Computers and Peripherals	www.ti.com/computers
Consumer Electronics	www.ti.com/consumer-apps
Energy and Lighting	www.ti.com/energy
Industrial	www.ti.com/industrial
Medical	www.ti.com/medical
Security	www.ti.com/security
Space, Avionics and Defense	www.ti.com/space-avionics-defense
Video and Imaging	www.ti.com/video

TI E2E Community

e2e.ti.com