White Paper

Using 3D Depth Sensing for Advanced Vision Applications

Abstract

3D time-of-flight (ToF) technology enables vision paired with high-resolution and high-accuracy depth perception for a wide range of applications, including robotics, machine vision, autonomous navigation, and augmented reality (AR). However, the path from depth data to application-specific algorithms is not straightforward due to the paucity of 3D-image–processing algorithms. Research is ongoing to develop efficient algorithms that combine artificial intelligence and 3D image-processing techniques.

In this paper, we explore 3D-image–processing solutions from Analog Devices Inc. (ADI) in addition to the existing 3D ToF software stack and development platforms, including state-of-the-art algorithms built on AI perception technologies. Application-specific solutions are presented, including people detection and tracking and volumetric measurement, along with descriptions of implementations and performance figures on different computing platforms and processors.

Introduction

Compared with traditional 2D cameras, which generate a color image of the scene, a 3D camera outputs a depth map in which each pixel represents the distance of the corresponding point in the scene to the camera. It is this depth map that enables enhanced object detection, classification, and tracking in the metric space.

ADI is a market leader in the 3D ToF space, providing complete solutions for high-resolution depth sensing, including ToF signal processors, analog front ends, laser drivers, and power products. The hardware offering is accompanied by open-source software components, enabling easy integration into Linux-based embedded systems as well as

AHEAD OF WHAT'S POSSIBLE™

Contents

interfacing with host computers for data acquisition and system control.

3D ToF depth-sensing systems are highly complex in terms of both hardware and software designs. To achieve competitive product development cycles, designers need solutions based on industry-standard hardware and software frameworks for faster system integration and development. Built on an industry-standard 96Boards[1] form factor, the AD-96TOF1-EBZ development platform[2] is based on ADI's ADDI9036 CCD signal processor and provides full system-level design, including a hardware platform, a complete software development kit (SDK),[3] additional collateral, and application examples.

The hardware design can be easily built into proofs of concept for a wide range of applications. Depending on individual preferences and development experience, various 96Boards processor boards as well as industry-standard development platforms, such as Raspberry Pi Nvidia Jetson or Xavier and NXP i.MX8, can be used for overall system evaluation and custom development. The system has options for MIPI, USB, and Ethernet connectivity to cover the full product development cycle, from evaluation, prototyping, and algorithm development on a host computer to full application development on an embedded system.

The open-source software SDK provided with the platform can be used throughout for a consistent software development experience and software reuse and stability all the way from the initial prototype to the final product. Windows and Linux support are built into the SDK as well as sample code and wrappers for various software frameworks and languages, including OpenCV, Python, MATLAB®, Open3D, and ROS.

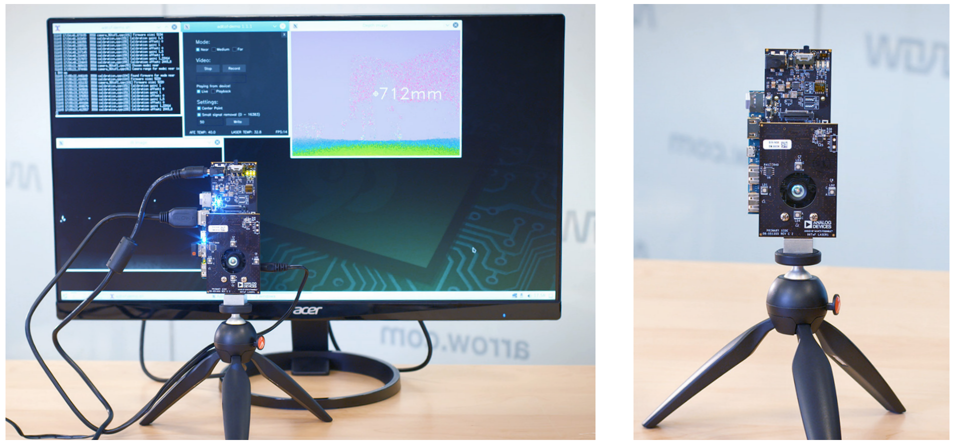
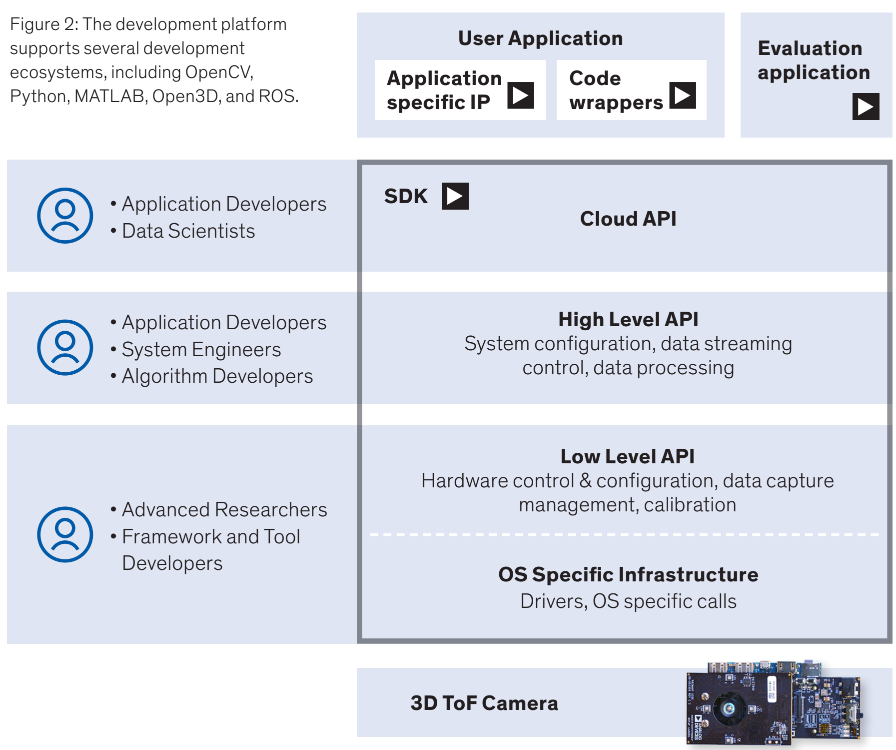


Figure 1: The AD-96TOF1-EBZ covers all product development stages, from evaluation to final application development on an embedded system.

Figure 2: The development platform supports several development ecosystems, including OpenCV, Python, MATLAB, Open3D, and ROS.



Use Case 1: People Detection & Tracking

Vision, for both humans and machines, is about detecting and recognizing objects in order to react to the surrounding environment. Recognizing people, measuring their position relative to the camera and other objects in the scene, and tracking their movements has many applications, including automatic door control, occupancy sensing, activity sensing, guiding collaborative and service robots, and collateral applications like contagion-spread prevention.

In automatic door control, 3D cameras can help prevent door damage by enabling the door-control system to react in time when people are moving toward the door with tall objects, like pipes and timber for construction. This is something that traditional sensors cannot “see.” By enabling precise people position measurement, 3D cameras help calculate trajectory and allow for the door-control system to keep the door closed when people are moving away or parallel to the door. This helps minimize unnecessary door opening, leading to HVAC cost reduction and door-life extension.

In the context of contagion containment, such as the Covid-19 pandemic, 3D-imaging technology can detect people, track their movements, and sense their position to help monitor both the number of people in a room as well as the distances between them.

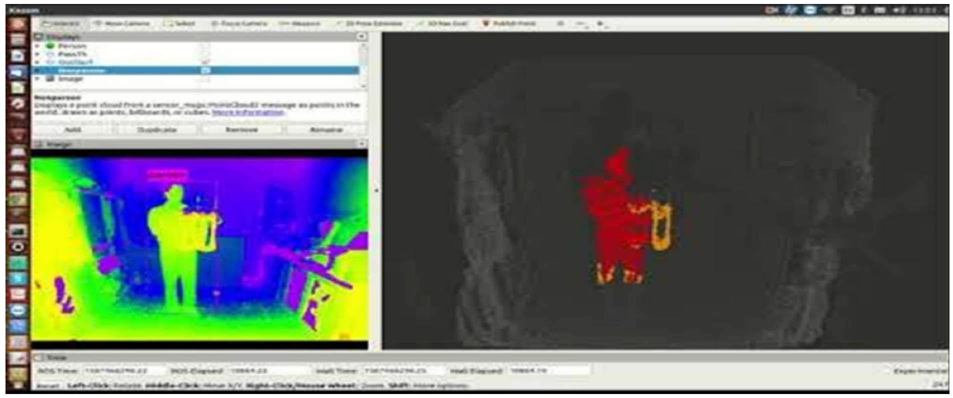


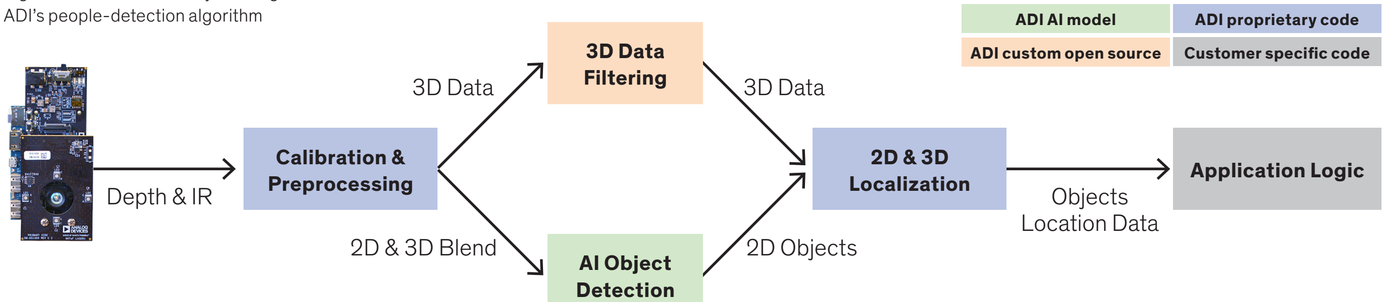
Figure 4: Running the algorithm in ROS on the Nvidia Jetson Nano to detect a person with an object and generate a connectivity list

Alternatively, it can simply count people going in and out a door to manage their numbers in the same space.

Similarly, reliably detecting people and their position is a must for collaborative and service robots as they become a bigger part of our lives at work, at home, and in public places.[4]

For such use cases, ADI provides efficient people detection, tracking, and counting software solutions that can be employed alongside ADI’s 3D ToF hardware. The algorithms are based on AI models that deliver enhanced detection rates in a wide number of scenarios and environmental conditions (Figure 3).

Figure 3: Data flow and key building blocks of ADI’s people-detection algorithm



A ToF camera, like any other type of camera, uses a lens to focus the light onto the sensor and employs one or more IR lasers to illuminate the scene. The field of view of the lens and the number of lasers and their field of illumination determine the ToF camera's sensing range, while the lens characteristics determine how light is distorted as it is projected onto the sensor. The data from the ToF camera therefore needs to go through a calibration and pre-processing stage that eliminates the effects introduced by the camera's specific optical configuration and ensures a high-accuracy and distortion-free depth image for enhanced point-cloud quality.

The pre-processed 2D and 3D data is then fed into an AI model for object detection and classification. The 2D data is the primary input into the AI model, while the 3D data is used to refine detection by adding detail to the objects of interest.

Platform	System Load	Frame Rate	People-Detection Rate
i7, four cores, 2.0 GHz	50% CPU 2 GB RAM	10 fps	85%
Nvidia Jetson Nano	80% CPU 70% GPU 2 GB RAM	12 fps	85%

Table 1: Detection-rate performance with various processors and frame rates

Like with any AI model, the detection rate and reliability are determined by the quality, size, and variety of the training data, and it is impossible to cover all situations with a given dataset. Therefore, the provided AI model can be improved with retraining, learning transfer, or feature extraction to better capture specific use cases.

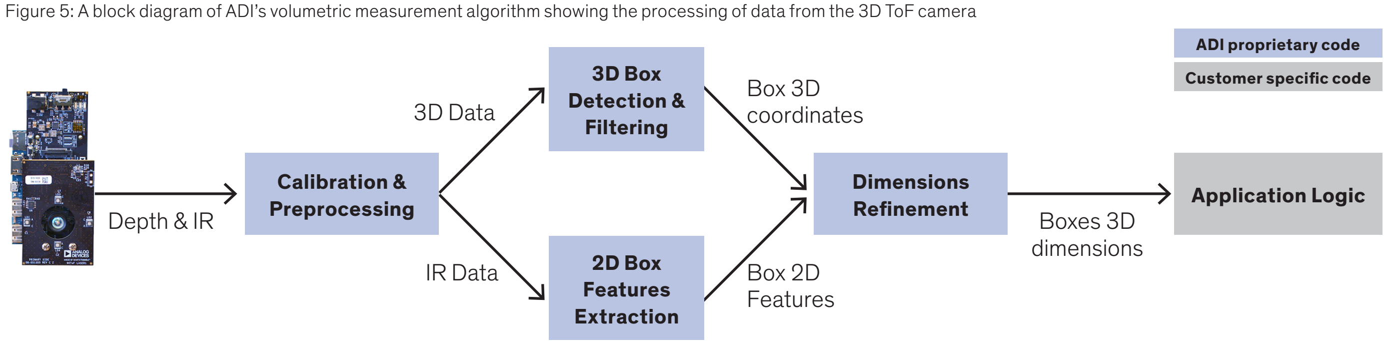
After the people and objects of interest have been detected by the AI model, they are localized by mapping them onto the depth image that has undergone filtering to reduce spurious noise as well as noise due to edge effects and reflections. The objects connected to a person are identified and added to a connectivity list for each person in the image (Figure 4). The 2D and 3D objects are then passed onto the application layer, which implements the application-specific logic that may control a door or a robot or generate warnings when social-distancing rules are not met.

Use Case 2: Volumetric Measurement

Humans can visually estimate distances and object sizes based on previous experience, but so can machines and devices equipped with 3D vision. The ability to accurately measure the size of a box regardless of its orientation and relative position to the camera is useful in logistics and industrial applications, such as cargo-space optimization for shipping companies, space optimization for transport and in storage businesses and warehousing, box sorting for industrial applications, box shape and dimension check for quality control, palletizing robots, and box fill-level sensing for automated packing lines. Automating box measurement in these applications can translate into millions of dollars saved annually by, for example, transporting full loads to reduce the number of trips and maximizing storage space in warehouses.

Accurate box measurement, however, comes with a set of challenges, as the relative position and orientation of the camera to the object of interest varies along with environmental conditions, such as material reflectivity,

Figure 5: A block diagram of ADI's volumetric measurement algorithm showing the processing of data from the 3D ToF camera



ambient light conditions, and the presence or absence of other objects on the scene.

As in the case of the people-detection application, the first step is to process the raw data to compensate for the effects introduced by the lens and calibrate the 2D and 3D data to account for the system-specific characteristics. This step also includes 3D-image undistortion, point-cloud generation, and point-cloud filtering to reduce noise and 3D artifacts (Figure 5).

Then the processing flow is split into two parallel paths, each working separately with 2D and 3D data. The 3D data path detects 3D surfaces to reconstruct the boxes in the scene. An initial box estimation is performed and the box edges refined to increase measurement accuracy.

The 2D data is used to detect the box edges and the relative orientation to the camera. Advanced filtering algorithms are applied for enhanced edge detection. Once the box information is available in 2D and 3D domains, the data from the two domains are correlated, and additional rotations and fitting are performed to improve 3D-dimension estimations. Finally, a list of boxes and their dimensions are passed on to the application layer to implement the specific use cases (Figure 6).

The frame rate and, consequently, the CPU load are determined by the scaling factors used for the depth map and the 2D image, which directly translates into the number of 3D and 2D points that will be processed. When running on an Nvidia Xavier AGX, the frame rate can vary between 5 and 10 fps and the processor load between 40% and 90%, depending on the chosen scaling factor.

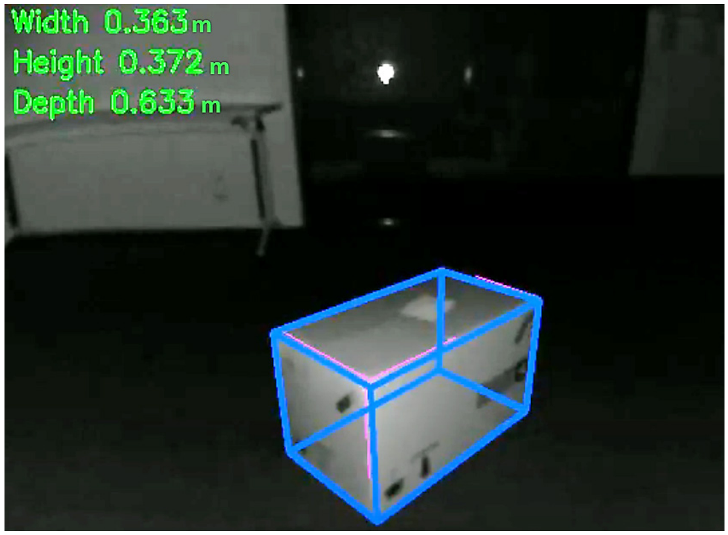
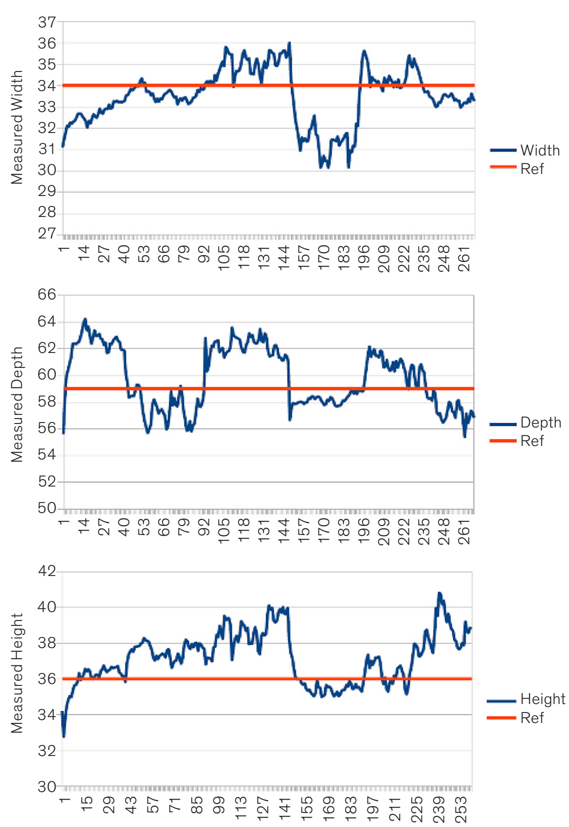


Figure 6: A box-detection and -measurement example running on the Nvidia Xavier AGX in the ROS environment

Figure 7: Experimental results showing the box-dimensioning performance



Summary

ADI's 3D ToF technology is an enabler of vision applications that require high resolution and high accuracy depth sensing. The AD-96TOF1-EBZ 3D ToF development platform is a comprehensive product solution designed to help engineers shorten product development cycles, and reduce risk and development costs through all the product development stages from system evaluation and prototyping all the way to production. By including examples and licensable algorithms for applications, such as gesture recognition, people detection and tracking, safety curtain for robots, and object detection for autonomous guided vehicles, ADI's 3D ToF platform eases final application development and provides customers with the confidence that the system is a perfect fit their specific application.

For more information about ADI's ToF technology, algorithms and available engineering services, visit https://www.arrow.com/tofalgorithms.

References

[1] 96Boards, https://www.96boards.org

[2] AD-96TOF1-EBZ ToF development platform, https://www.analog.com/en/design-center/evaluation-hardware-and-software/evaluation-boards-kits/ad-96tof1-ebz.html

[3] Analog Devices ToF SDK, https://github.com/analogdevicesinc/aditof_sdk

[4] Service Robotics Market — Global Forecast to 2025, https://www.marketsandmarkets.com/Market-Reports/service-robotics-market-681.html

Five Years Out

Arrow Electronics, Inc.
7459 South Lima Street
Englewood, CO 80112, USA