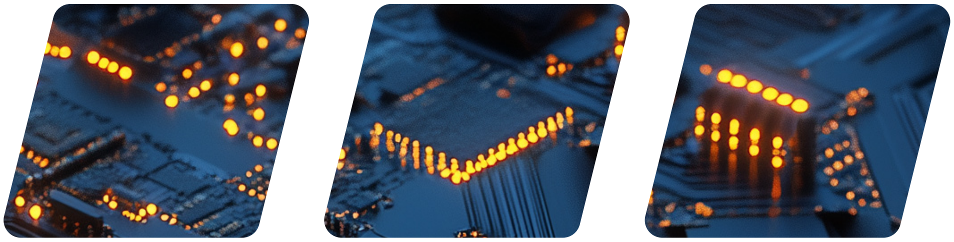
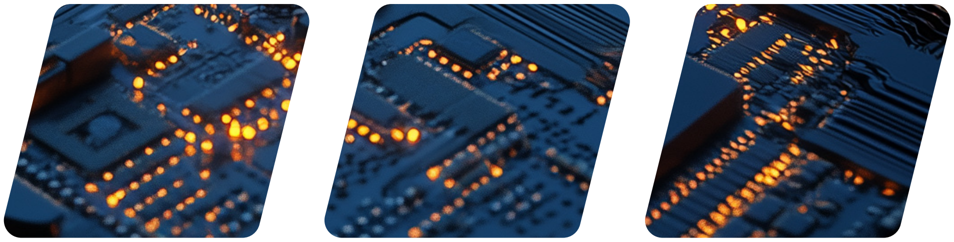


Random Bin Picking Based On Structured-Light 3D Scanning



White Paper

Author:
Hoon Choi, Sr. Fellow, R&D, Lattice Semiconductor

DISCLAIMERS

Lattice makes no warranty, representation, or guarantee regarding the accuracy of information contained in this document or the suitability of its products for any particular purpose. All information herein is provided AS IS, with all faults, and all associated risk is the responsibility entirely of the Buyer. The information provided herein is for informational purposes only and may contain technical inaccuracies or omissions, and may be otherwise rendered inaccurate for many reasons, and Lattice assumes no obligation to update or otherwise correct or revise this information. Products sold by Lattice have been subject to limited testing and it is the Buyer's responsibility to independently determine the suitability of any products and to test and verify the same. Lattice products and services are not designed, manufactured, or tested for use in life or safety critical systems, hazardous environments, or any other environments requiring fail-safe performance, including any application in which the failure of the product or service could lead to death, personal injury, severe property damage or environmental harm (collectively, "high-risk uses"). Further, buyer must take prudent steps to protect against product and service failures, including providing appropriate redundancies, fail-safe features, and/or shut-down mechanisms. Lattice expressly disclaims any express or implied warranty of fitness of the products or services for high-risk uses. The information provided in this document is proprietary to Lattice Semiconductor, and Lattice reserves the right to make any changes to the information in this document or to any products at any time without notice.

INCLUSIVE LANGUAGE

This document was created consistent with Lattice Semiconductor's inclusive language policy. In some cases, the language in underlying tools and other items may not yet have been updated. Please refer to Lattice's inclusive language FAQ 6878 for a cross reference of terms. Note in some cases such as register names and state names it has been necessary to continue to utilize older terminology for compatibility.

ABSTRACT

The random bin picking using the structured-light 3D scanning plays an important role in factory automation. It enables the robot to pick-up known objects with random poses out of a bin and sort out the objects for the next assembly step. This white paper describes how the Lattice FPGA benefits the system design, specifically by reducing the system BOM cost. By aligning with Lattice's FPGA solutions, you can build a cost-effective random bin picking system.

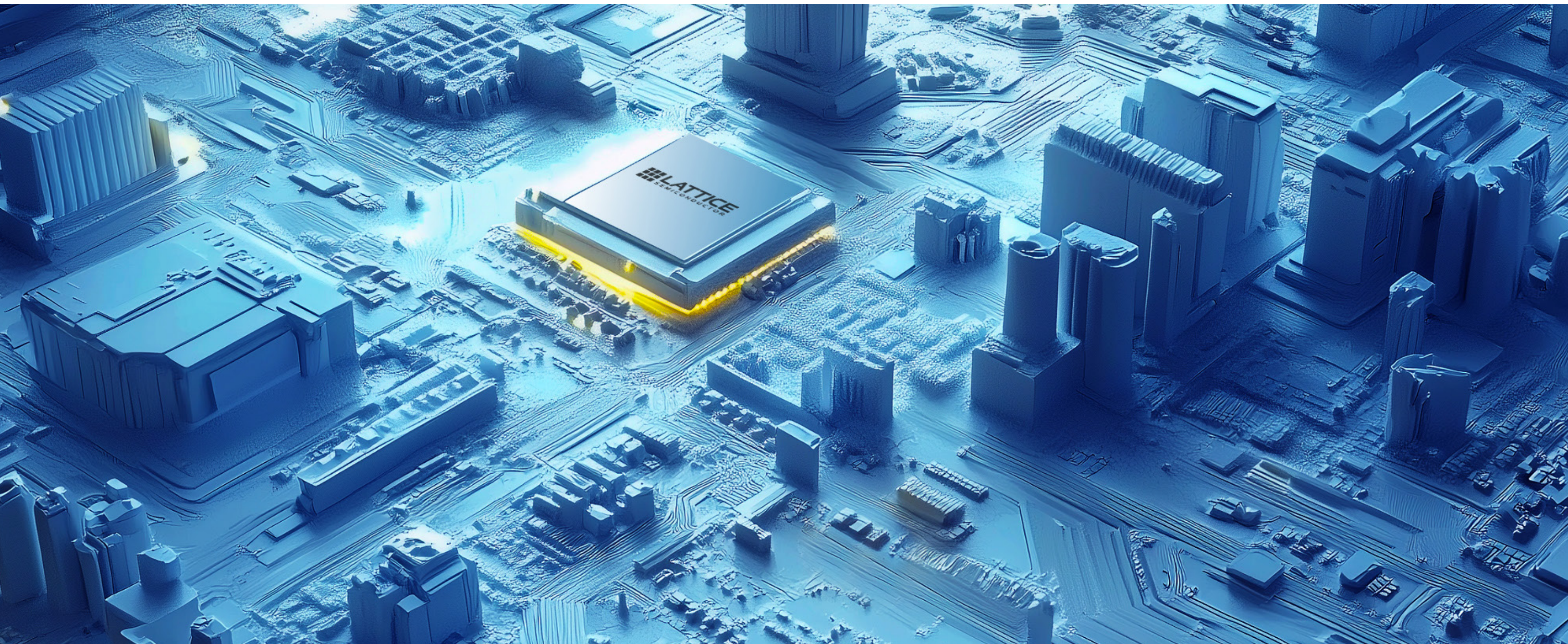


TABLE OF CONTENTS

1. Background of Structured-Light 3D Scanning

Structured-light 3D scanning is used to capture the three-dimensional shape of an object by projecting light patterns, such as grids or stripes, onto its surface. The deformation of these patterns is recorded by cameras and processed using specialized algorithms to generate a detailed 3D model. See Figure 1.

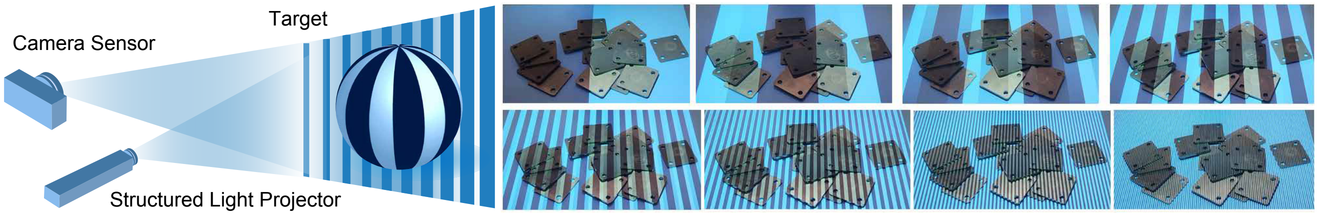
It is widely employed in fields such as industrial design, quality control, augmented reality gaming, and medical imaging. Compared to laser-based 3D scanning, structured-light scanners use non-coherent light sources, such as LEDs or projectors, which enable faster data acquisition and eliminate potential safety concerns associated with lasers. However, the accuracy of structured-light scanning can be influenced by external factors, including ambient lighting conditions and the reflective properties of the scanned object.

Projecting a narrow band of light onto a three-dimensional surface creates a line of illumination that appears distorted when viewed from perspectives other than that of the projector. This distortion can be analyzed to reconstruct the geometry of the surface, a technique known as light sectioning.

A more efficient and versatile approach involves projecting patterns composed of multiple stripes or arbitrary fringes simultaneously. This method enables the acquisition of numerous data points at once, significantly improving scanning speed. When viewed from different angles, the projected pattern appears geometrically distorted due to the object's surface shape, allowing for precise surface reconstruction.

While various structured-light projection techniques exist, parallel stripe patterns are among the most used. By analyzing the displacement of these stripes, the three-dimensional coordinates of surface details can be accurately determined.

Figure 1: Structured-Light 3D Scanning



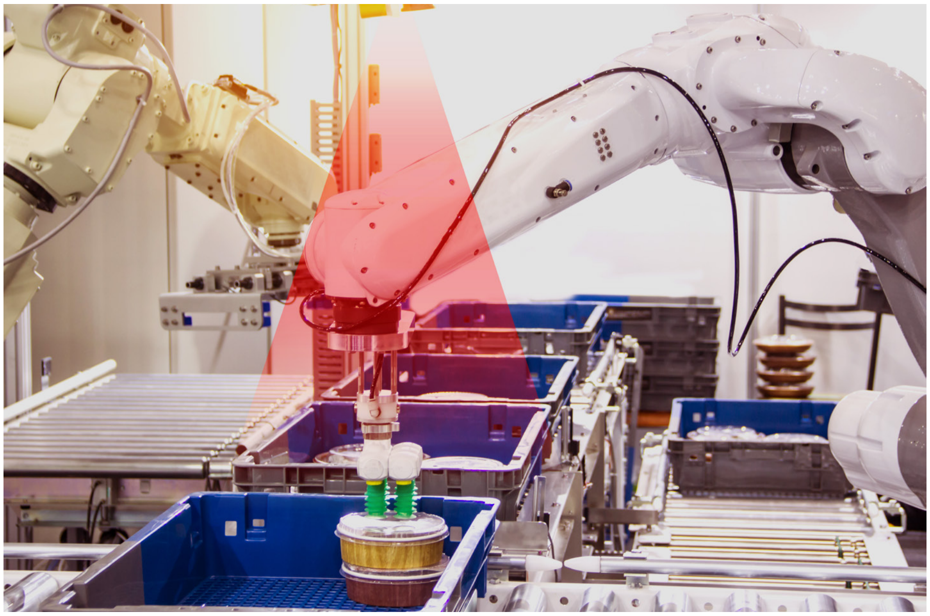
2. Background of Random Bin Picking

Random bin picking is a core problem in computer vision and robotics. The goal is to have a robot with sensors and cameras attached to it pick up known objects with random poses out of a bin using a suction gripper, parallel gripper, or other kind of robot end effectors. This random bin picking is mainly used in factories to sort out objects for the next assembly step from the bin of randomly stacked objects.

Modern random bin picking usually uses structured-light 3D scanning techniques to get the detailed 3D information of each object. From the 3D information, it recovers the shapes of objects and determines their orientations in space. Computer vision plays the main role in object identification and picking points calculation. For complex objects, CAD information (e.g., detailed shape information) given by users is also used to facilitate the object detection by geometrical matching. Recently, machine learning-based approaches such as object detection and segmentation are employed to handle more complicated cases. One important factor in the calculation is to identify the overlap of objects so that the robot arm picks the one that is not under any other object. Otherwise, due to the other objects over the target object, the picking may fail. Based on the calculation, robot arm picks and moves one object to the target location.

The solution is generally composed of two modules that are connected through Ethernet: the sensor module and the computing module. The sensor module projects the structured light to the bin and captures the reflected image using a camera. The captured images go to the computing module that has a powerful computing resource (e.g., CPU, GPU) to do all the computing mentioned above. The final target coordinate is informed to the robot control module. See Figure 2.

Figure 2: A Sensor Module Projects Structured Light to the Bin



3. The Lattice Solution

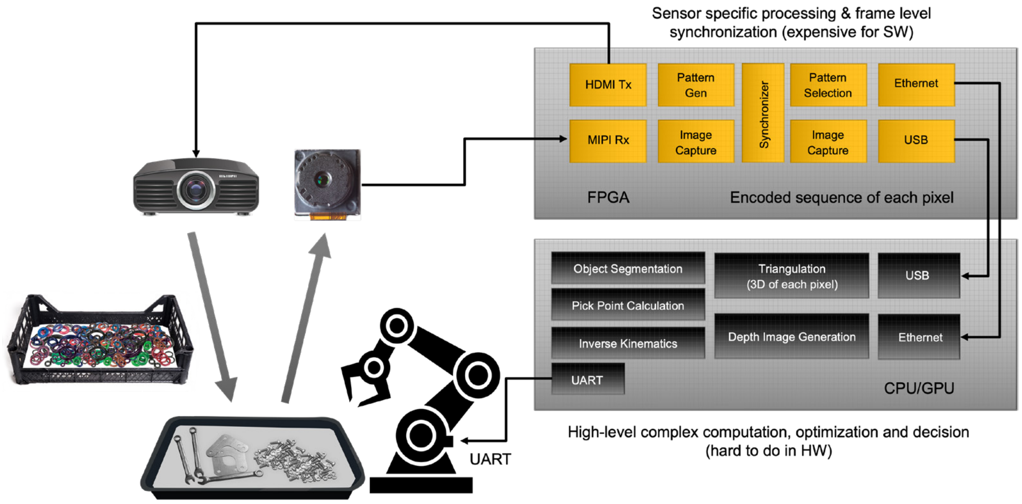
In our solution, we partition the problem in such a way that the FPGA in the sensor module offloads the computing tasks for the computing module and reduces the bandwidth of the Ethernet connection between the two modules by sending one encoded image instead of a sequence of raw images. This speeds up the sensing part and also reduces the BOM of the computing module.

In the sensor module, for the structured-light 3D scanning, the FPGA generates the sequence of images and sends one image at a time to a projector. After sending one image, the FPGA triggers a camera sensor to capture the bin image that corresponds to the image just sent to the projector. This guarantees the frame-by-frame synchronization and enables the change of image at every frame for a fast scanning.

The gray binary pattern we use in the generated images includes positive, negative, horizontal, and vertical patterns as well as all white and all black. The number of images is 41. The sequence of captured images at the camera comes back to the FPGA over the MIPI CSI link. The FPGA encodes the images into a 10-bit coded image that identifies the location of the corresponding pixel in the generated images. This coded image is passed to the computing module. This encoding reduces the bandwidth of the Ethernet link to the computing module significantly. For example, for a 1080p case, we need to send 1920 x 1080 x 8b x 41 = 680 MB of data if we send a raw image, while it reduces to 1920 x 1080 x 10b x 2 = 41MB (16X reduction) for an encoded image. This significantly reduces the bandwidth requirement of the Ethernet link.

In the computing module, triangulation generates the depth image from the given encoded image. This depth image is used for object detection (segmentation) and pick point calculation. See Figure 3.

Figure 3: The Lattice Solution

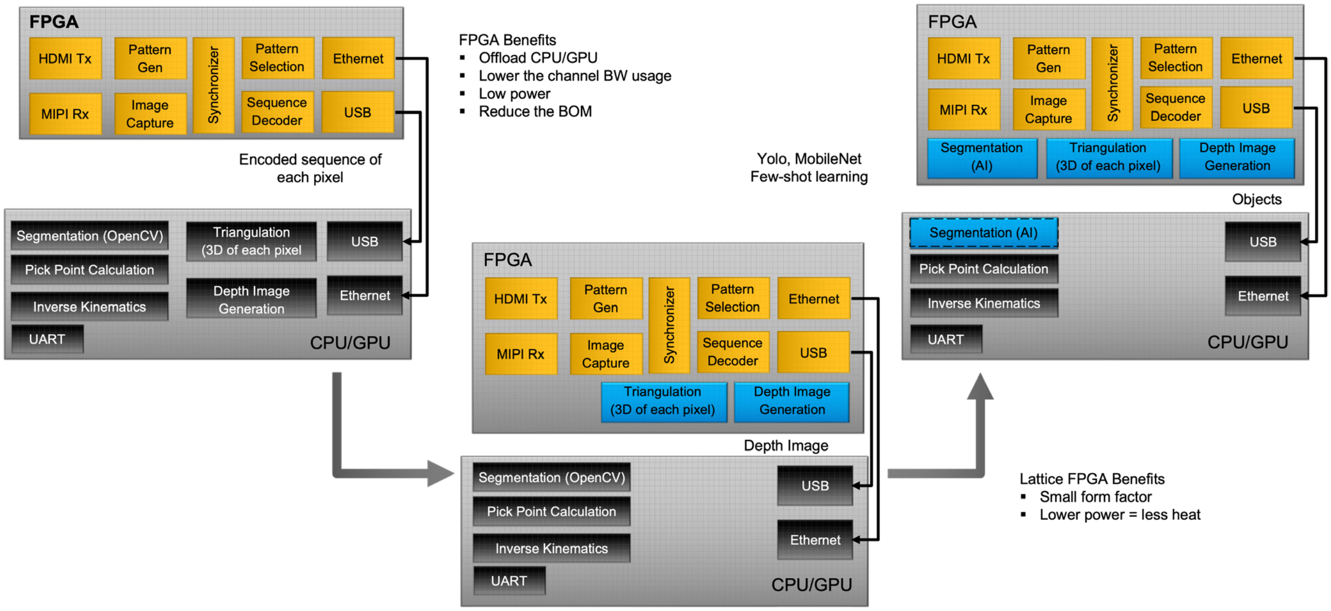


The triangulation needs to be done for each pixel to get the depth information of each pixel. This is very repetitive and pixel-level computing that can be done in parallel. Lattice FPGAs can offload this task from the computing module by doing the triangulation and generating the depth map. The generated depth map can be transferred to the computing module. This reduces the computing requirements of the computing module, hence the user can use lower-end computing resources and reduce the BOM. Alternatively, the user can use the same computing resource but add extra capabilities since triangulation part is moved to FPGA. Similarly, the FPGA can perform all or part of machine learning-based object detection and segmentation to offload the computing module further. See Figure 4.

The target FPGA for the task up to encoded image generation is Lattice CrossLink^{TM}-NX or Lattice Certus^{TM}-NX, while Lattice CertusPro^{TM}-NX, Lattice Avant^{TM}, or Lattice Certus^{TM}-N2 (Lattice Nexus^{TM} 2) is needed for further offloading. External memory like HyperRAM or LPDDR is needed to store the captured images and generate the encoded image. LPDDR is needed to do further offloading with the generated depth image.

The small form factor and the low power consumption of Lattice FPGAs enable the small-sized and low-cost sensor module since it can fit inside of a plastic enclosure without any extra components for power dissipation (such as a fan, heatsink, etc.).

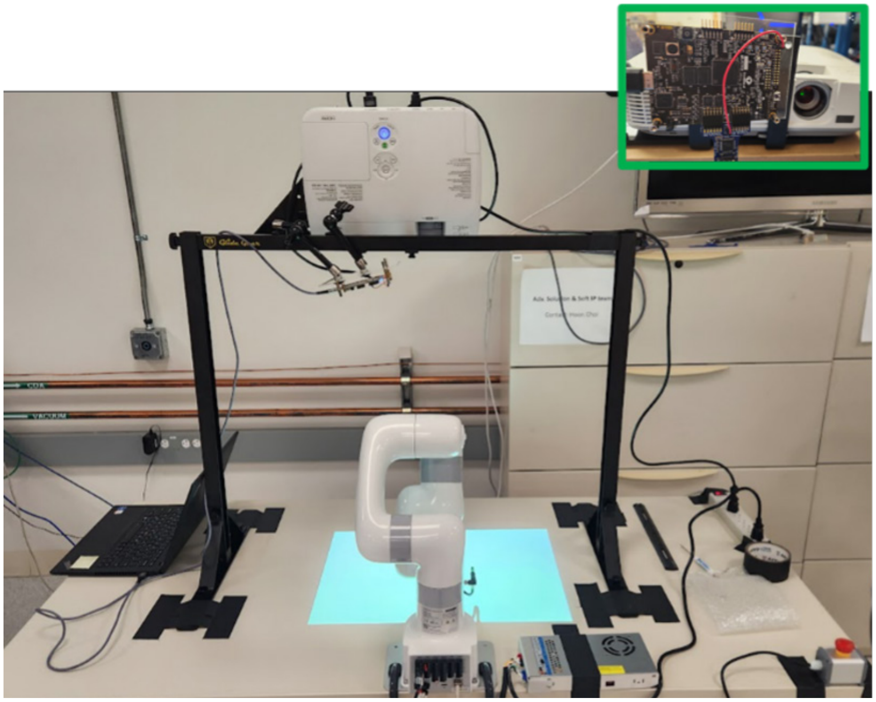
Figure 4: Lattice FPGA Benefits



As the development platform to verify the concepts and algorithms, a proof-of-concept demo system was built. The current proof-of-concept demo was built with a general-purpose projector (NEC NP-M300X LCD projector) and CPNX VVML development board. The computing module is an NVIDIA Jetson Orin Nano. For the robot arm, we used UFACTORY LITE6 that is controlled by Python code.

The use of a general-purpose projector limits some of the capabilities. For example, the projector performs frame-by-frame image compensation to make it clearer to human eyes. However, it impacts the pattern shape and brightness, thus we need to wait until it goes away, resulting in every other frame image change at the max instead of every frame. In addition, the native input resolution of the projector is XGA (1024 x 768), and if we feed a 1080p image, the scaler in the projector dithers and breaks the patterns. So, in the demo, XGA is used. For products, DLP-based fast frame rate projectors are recommended. See Figure 5.

Figure 5: Proof-of-Concept Demo System



Link to video: https://www.youtube.com/watch?v=2FVouhpYkL8

4. Conclusion

Random bin picking using structured-light 3D scanning is supported by Lattice FPGAs. The use of Lattice FPGAs enables the reduction of the BOM in two modules: 1) in the computing module due to offloading to the FPGA, and 2) in the sensor module due to the lack of necessity for heat dissipation components in the plastic enclosure.

READY TO LEARN MORE?

To learn more about Lattice low power FPGA-based solutions for industrial, automotive, communications, computing, and consumer applications, visit **www.latticesemi.com** or contact us at **www.latticesemi.com/contact** or **www.latticesemi.com/buy**.

TECHNICAL SUPPORT ASSISTANCE

Submit a technical support case through **www.latticesemi.com/techsupport**.
For frequently asked questions, please refer to the Lattice Answer Database at **www.latticesemi.com/Support/AnswerDatabase.**

V.1 WP0043