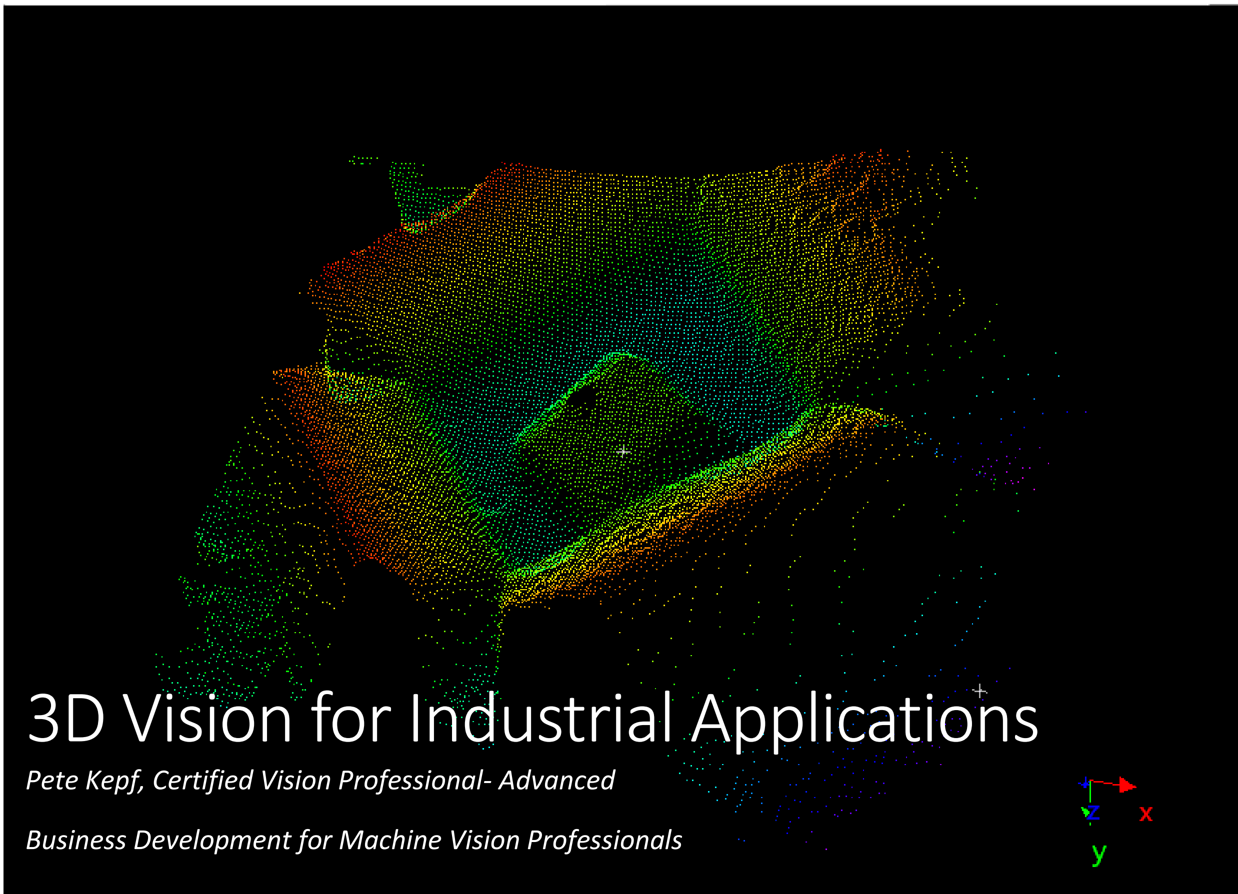


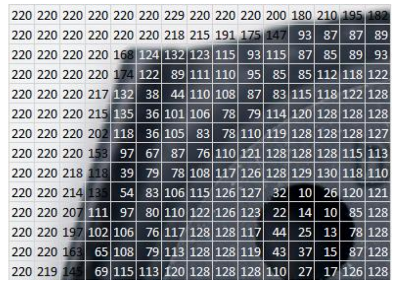
Introduction

Since its inception, machine vision professionals have used a variety of creative techniques, with varying degrees of success, to attempt to extract 3-dimensional information from 2D images. Barring that, 3D sensors were incorporated into traditional inspection systems to obtain critical depth information. Today, 3D vision is coming to its own with an arsenal of technologies to support a growing number of 3D vision applications. In the early days of 2D vision, the technology was evolving so rapidly and that the application expertise lagged the systems' capabilities. So today with respect to 3D vision, the potential application demand exceeds the experience base to deliver robust solutions. We're still learning.

And as the increase in technology capabilities continues to produce a migration of 2D applications to those once confined to the factory floor, a similar circumstance can be found in 3D vision. Recently, Vision Systems Design highlighted an application to monitor the behavior of pigs (Vision Systems Design, January 2019). Scholarly articles discuss a wide variety of 2D and 3D vision applications: in agriculture (Lyndon Neal Smith*, 2018); in construction to monitor workers (Yong-JuLeeMan-WooPark, 2019) and of course, Apple iPhone face ID. But the application base for 3D vision on the factory floor is also exploding, so for the foreseeable future the industrial machine vision professional can satisfy his natural curiosity and contribute to the experience with such applications as inspecting shiny parts (Andrew Wilson, January 2019), determining the location in space of fine wires (Eckhard, January 2019), robotic bin picking (https://www.photoneo.com/bin-picking/) and general 3D part inspection techniques. Here we'll review historically applied techniques, still valid today, as well as explore the latest technologies, remarkable for their price performance as well as their sheer capabilities.

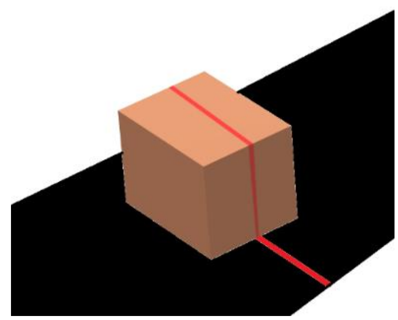
2D Depth Perception

A 2D vision system may be thought of as a device that obtains brightness information about points within a grid that is superimposed onto a scene or object. The point grid is the pixel matrix on the vision sensor and the brightness is the amount of light that is collected by each pixel (color sensors use filters to capture various amounts of light at certain wavelengths, but brightness is the parameter that is obtained). Software algorithms then calculate attributes about the object based on Width (x dimension), height (y dimension) and intensity level (brightness). Although the world is flat for 2D vision systems, there are some creative techniques that have been applied over the years to obtain 3D information from a 2D image.



Structured Light

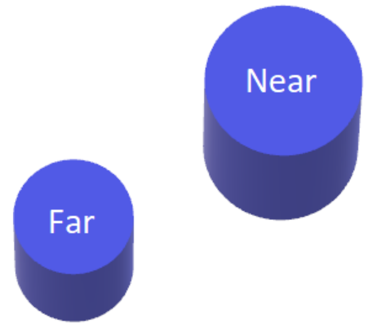
Perhaps the simplest method of obtaining 3D information from a 2D image is the use of structured light. If a line is projected over an object, the shape of the line's reflection will correspond to the shape of the object. The location of the line can be determined by the brightness of pixels within the grid that correspond to the light line. The positional displacement of the line then corresponds to, in the example at the right, the length and height of the box. This technique can be quite accurate and robust, especially when filters are applied (for example, red). However, in order to be accurate, the area of consideration must be in a fairly restrictive location and surface finish may also adversely affect the results.



Industrial machine vision applications include package size and shape, part position component verification and a variety of image enhancements.

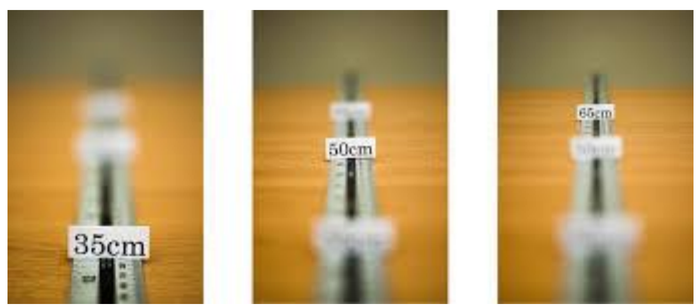
Relative Size

Another fairly simple method of obtaining 3D information with a 2D system is to measure size variation. Of course, the nearer an object is to the sensor, the larger it appears. Given a population of uniform objects the relative distance from the part to the camera can be determined by measuring the size of the object. The advantage to this method is that more than one object can be detected at a time. Disadvantages include that the objects must be presented in the same plane, accuracy can vary greatly with part geometry and object size must be known ahead of time.



Focus

A third way to determine 3D location information is to detect whether or not an object is in focus at a given distance. This method requires both a sharpness algorithm for the image and an encoder signal for position, either from the lens or from the camera stage. This method is commonly used in photography and outdoor range finding (see figure at right (Deepview, Enhancing Perception Through Eye-tracking, 2019), but lacks the precision required for manufacturing processes. Note that, the only information provided is depth; no other information about part geometry or other attributes is obtained.



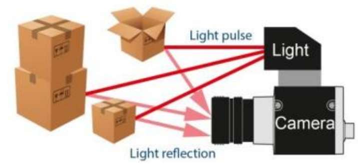
Hybrid systems of the above with standard machine vision have been applied quire successfully. For example, palletizing/ depalletizing operations that combine an optical, infrared or ultrasonic sensor to provide depth information to a linear actuator the moves a camera to the appropriate height for inspection are not uncommon. Similar techniques have also been applied to provide depth information to robots for pick and place operations. But these techniques lack the elegance of true 3D, which is emerging as the technology of choice for not just guidance, but also inspection.

3D vision

A 3D vision system adds depth (z dimension) to height and width. Intensity is still also considered, but is manipulated differently, as will be discussed below. Some 3D vision sensors have been around for years and others are quite new, but the real breakthroughs have come from the technological advances in manipulating large amounts of vision-based data. These advances, as well as price reductions as more sensors are developed and sold, have resulted in machine vision application engineer applying 3D vision systems where once only a 2D system would have been considered. Common sensor technologies include Time of Flight (TOF), stereo imaging, laser profiling and fringe projection.

Time of Flight (TOF)

A TOF camera provides depth by projecting a pulse of light into the scene and then measuring the phase shift between the pulse and the reflected light. This phase difference is a function of the distance between the reflector and the sensor. Usually the light pulse is in the IR range, invisible to humans and the receiving sensor is sensitive in the same range. Thus, X and Y are provided by the location within the pixel grid and Z is provided by the time taken to read the light signal (phase shift). The result is either a depth map, which depicts distances as either varying brightness or color, or a point cloud, which is a true 3D mesh representation.



Graphic courtesy Basler AG

Because the TOF camera provides its own illumination source, its influence by ambient light is low, except where IR may be present, as in sunlight. TOF cameras are also relatively compact with a fast scan rate. Disadvantages are lower resolution and higher cost than some 3D options.

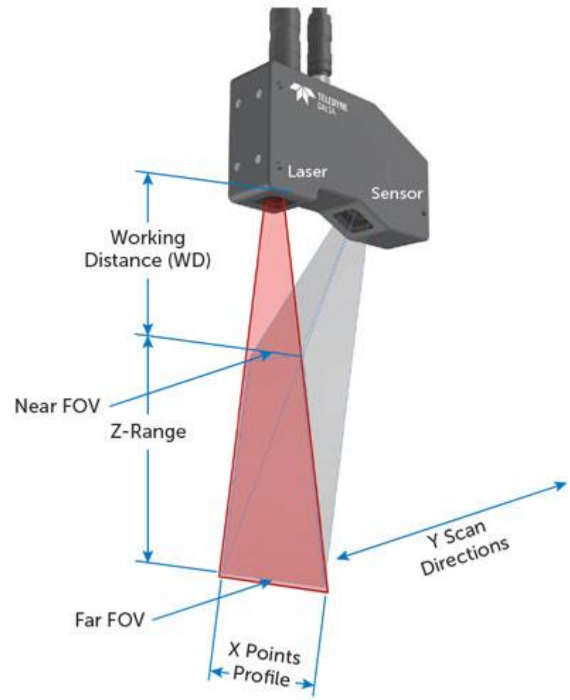
Stereo Imaging

Stereo Imaging involves the use of two cameras a given distance apart. These are used to acquire images of the same scene and the images are compared to one another. By triangulating the features in the two images based upon a known distance between the cameras, the distance to objects in the scene can be calculated. Stereo imaging works similarly to the human eyes, but without the high resolution, color, massively parallel processed imaging information coupled with the motion sensor (muscles).

Because neither camera lens/ imaging sensor is parallel to the scene, distortion is always present. This effect can be minimized, but not entirely eliminated by increasing the distance between the sensors and the objects during both calibration and operation. In addition, contrast across the scene is required to differentiate objects. For example, stereo imaging cannot tell the distance between a blank wall and the sensor because there is no contrast. In instances such as these, patterns of light have been projected onto the surfaces to provide contrast.

Laser Profiling

Laser profiling is a popular technique and mature, robust technology. The sensors are inexpensive and resolution is good. The laser profiler works similarly to structured light methods in that a laser line is projected onto an object and the reflection is captured by a linear imaging sensor. Displacement is calculated by triangulation, which provides the 3d information. Unlike the structured light methodology, however, the distance between the laser projector and the CMOS line sensor is precisely engineered and the profiler is calibrated for this distance.



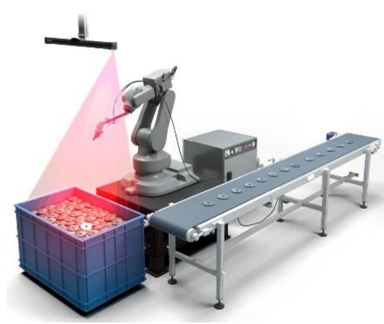
Graphic courtesy Teledyne Dalsa

Being linear, the laser illumination source provides information in one axis (x) only. Z axis information across a range of distances is calculated by triangulation. Y axis information is provided through an encoder input as the object moves across the scanner. Therefore, accuracies in x and z are determined by the scanner specifications and accuracy in y by the resolution of the encoder.

This technique is not applicable to stationary objects unless a method is employed to move the scanner (e.g. with a robot) relative to the scene. For moving parts, however, scan times are very fast and accurate.

Fringe Projection

Fringe projection, sometimes referred to as "pattern projection" is a method whereby a series of patterns of varying angles and geometries is projected onto the scene. Using a combination of triangulation and specialized optics and calibration algorithms, 3D information is extracted from the scene. The output of the 3D scanner is usually a point cloud which may be input to a variety of image processing environments. These software development packages can solve a host of problems from inspection of stationary parts to precision robot guidance.


Graphic courtesy Photoneo s.r.o.

As with laser profiling, the scanner consists of a line transmitter and a CMOS sensor receiver. In the case of Fringe Projection, however, the sensor captures 2D information as opposed to a line. Thus, x and y axis information is a function of the scanner design and the z axis information is calculated through triangulation.

This technique also employs sophisticated optics and calibration algorithms to provide a high degree of accuracy in all three dimensions. These algorithms also provide accurate scalability across a wide z dimension, making it ideal for high-precision, stationary part applications such as robot bin picking. In addition, the 2D image sensor capabilities allows the capture of standard images for 2D inspection.

3D Machine Vision Applications

Although methodologies vary, 3D machine vision applications comprise three broad categories: Mapping/ data input, Inspection and Guidance.

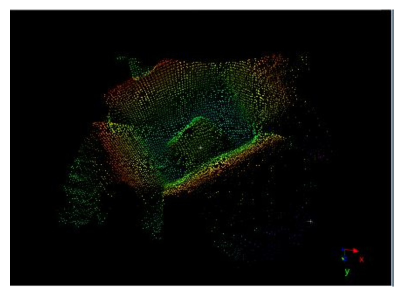
Mapping

Mapping is simply inputting z axis information for distance measurement or, more complexly, gathering point cloud information about a scene or object for use by another process. Range finding/ auto-focus is a simple mapping process use in Intelligent Traffic systems. Apple face ID and Microsoft's Kinect are examples of consumer-based technologies whose concepts are being applied in industry, An Autonomous Vehicle, for example, may have a pattern emitter which reflects from various objects and are then used by on-board processors to define a virtual work environment. Another mapping example is the gathering of point cloud information by a scanner for reverse engineering or 3D printing.

Photoneo Phollower 100

Inspection and Detection

As 3D vision systems become more affordable, machine vision application engineers are taking advantage of the additional dimensional data to perform part inspection. In some instances, 3D vision may be less susceptible to the vagaries of part finish, making it a preferable tool for dimensional checks. Inspection tasks such as part verification, assembly verification of shiny metal parts, inspection of low-contrast surfaces such as tires and package fill inspection are examples of 3D point cloud information from a scanner being input to a generic machine vision software package.



Container Point Cloud Profile

Some 3D scanners meet industry standards for robot safety zones. This allows a scanner to be positioned in the work cell and various approach zones to be defined to alter robot behavior based on the presence of a human.

Industrial Robot Guidance and Vision Guided Bin Picking

Solutions for industrial robot guidance and vision-guided bin picking are evolving due to the proliferation of lower cost Graphical Processing Units (GPUs). These processors, originally designed for the gaming industry, enable the manipulation of the large amounts of 3D point cloud data generated by the scanners. Coupled with the accuracy/ calibration features mentioned above, it is now practical to design robust solutions to guide industrial robots to pick up parts based upon their surface features or even by their geometry.

Random Bin Picking.

Photo courtesy Photoneo s.r.o.

Random Bin Picking

Random bin picking describes a collection of parts that is defined only by a portion of its exposed surface. For example, a tube of toothpaste, a deodorant dispenser and a shampoo tube may all have differing geometries, but may share an area that is similar among them in size and profile. A subset of this area of interest may then be defined as a pick point for a generic robot end effector (e.g., a suction cup). The vision software does not "know" what part it is picking up; only that it found a 3D surface profile that matches that which was taught. Neither does it necessarily "know" the part orientation. This method is useful for generic transfer of similar objects (cartons, tube, etc.) for packaging or process applications.

3D Bin Picking

3D bin picking applications are quite comprehensive in that they use 3D point cloud for all objects of consideration: the part, the end effector, the robot, and the work space. A 3D CAD file is used to generate a point cloud model of the part. Likewise, CAD data is used to build point cloud models of the remaining components. This information is then combined into a virtual 3D workspace. With this information, precise part pick points can be defined, as well as gripper offsets, robot optimization paths and collision avoidance.

3D Bin Picking

Photo courtesy Photoneo s.r.o.

Summary

3D vision applications are becoming more and more popular as technological capabilities increase and costs decrease. These factors, combined with an expanding of the application base as engineers become more familiar with the capabilities and limitations of 3D vision imply that this technology, still in its infancy, will continue to grow.

Pete Kepf, CVP-A provides Business Development for machine vision providers.

References

Andrew Wilson, E. E. (January 2019). Novel inspection system characterizes shiny reflective parts. *Vision Systems Design*, 12-13.

Deepview, Enhancing Perception Through Eye-tracking. (2019). Retrieved from Project Overview: https://deepview.cs.st-andrews.ac.uk/overview/

Eckhard, T. (January 2019). Using 3D stereo vision to inspect bonding wire. *Vision Systems Design*, 14-16.

Lyndon Neal Smith*, W. Z. (2018). *Elsivier.com.* Retrieved from Computers in Industry 901: http://eprints.uwe.ac.uk/35364/7/Innovative-3D-and-2D-machine-vision-methods-for-analysis-_2018_Computers-in-.pdf

Vision Systems Design. (January 2019). 3D vision system monitors the behaviour of pigs. *Vision Systems Design*, 5-6.

Yong-JuLeeMan-WooPark. (2019, February). *El SivierL Automation in Construction.* Retrieved from 3D tracking of multiple onsite workers based on stereo vision: https://www.sciencedirect.com/science/article/pii/S0926580518308379