TECH BRIEF

Which method of 3D imaging is best for my application?

Each method has merits:

- Triangulation
- Structured light
- Time of Flight (ToF)
- Stereo vision

Introduction:

Human vision is binocular (two-eyed) and in color, and typically serves us well for a wide range of activities. We differentiate objects, sense small and large objects, gauge depth and distance, and so on. We've adapted mechanical tools like measuring sticks and calipers, or electronic tools like radar, to augment our vision, and to help us navigate and conduct survival skills in real-time.

Machine vision, on the other hand, is about automating vision tasks, often special purpose tasks, and often to very high levels of precision. The means by which machine vision is accomplished need not, and often do not, draw upon the same methods used for human vision. If a machine vision method like stereo vision correlates somewhat to human vision, that may be interesting to a biologist or a cognitive scientist, or a vision systems designer, but it's beyond the scope of this Tech Brief.

Our objective it to survey the key methods of 3D imaging for which mature products are available, to help the reader understand the range of options. Whether you plan to configure and optimize your vision system by integrating components yourself, or to instead engage an integrator or purchase a turnkey system, it can be helpful to survey the different methods available for 3D imaging, and their key differentiators.

3D Imaging (vs. 2D vs. 1D)

Do I need 3D to solve my problem? Excellent question. It's a variation on the general principle of avoiding over-engineering, controlling costs, and prioritizing the problem solving outcomes as the key point. Let the method of the solution become only as complicated as necessary to reliably solve the problem.

Of course one may need to build in a bit of margin to handle outlier situations, but if a problem may be constrained to controlled conditions, it can simplify the engineering design, the implementation, and the costs, while achieving very effective outcomes.

Unsurprisingly, the history of machine vision evolution has built from simple early solutions in 1D, through 2D, and into 3D, in more or less direct correlation with Moore's Law. As electronics, on both the sensor-side and the compute-side, have become more powerful and less expensive at the same time, systems designers have exploited the evolving capacity to automate ever more complex solutions.

1D machine vision

As a trivial example, consider the decades-old "electric eye". Whether using infrared or visible light, a typical application might open a door when a leg breaks the beam, or stop the lowering of an overhead garage door if a car body or other object blocks the light. It's not very sophisticated, but it's practical.

A more complex 1D application is a bar code reader – the kind used for grocery kiosks. The varied width lines in the bar code are scanned with a linear light source and the reflection is read by a narrow sensor and interpreted by an algorithm that can cope with a certain amount of angular deviation from the optimal 90 degree orientation.

2D machine vision

Sticking with code reading, QR codes can encode far more information than a bar code, but requires a two-dimensional sensor to digitize the code.

Counting non-occluded objects in a bin, or identifying cell clusters on a microsope slide, are also 2D applications.

In short, if one only needs the X and Y dimensions to create the solution, and there is no Z dimension or it is irrelevant, it's a 2D machine vision application. Several decades of machine vision growth are in the 2D space, which continues to expand with new innovations.

3D machine vision

There are a number of applications for which one needs the third dimension, Z. In some cases it's exactly a degree of precision for Z that is the key value being determined. Example: If you are doing an automated milling or grinding application, for which the process continues until a certain (narrow) tolerance is achieved, the precision of the Z value may be expressed in values of just a few microns.

In other cases one needs X and Y with high precision, and a Z value that's "good enough". Suppose a robot arm is to reach into a bin of jumbled parts of random (X,Y) orientation. Further parts may be flat on the bin floor, leaning against each other at angles, or on top of each other. The rubberized robot pincher may be adept at picking items gently from any (X,Y) orientation, but needs enough Z data to determine whether to "go deep" to an object on the bin floor, or "remain high" to an object stacked 3 or 4 layers up. If the parts in question are 10cm on a side or more, one doesn't need micron-level data on the Z dimension – precision to within a few millimeters is good enough.

Keep that in mind as we dig in below – optimal outcomes need theoretical underpinnings – but the best outcome for a given application is always relative to the objective one is trying to achieve.

Key Concepts:

So your application requires a 3D approach – how does one proceed? Is there a standard way to do 3D machine vision, or are there multiple ways to approach it?

Initial 3D systems were mostly based on triangulation – an effective approach for many applications that remains popular and continues to evolve today. But as with many fields of engineering, innovation has led to a number of different approaches to 3D imaging, each of which has compelling rationale – under the right conditions.

Let’s name the four current approaches to 3D imaging before we go into more detail: (numbering does NOT imply ranking or quality – each has merits!)

1. Triangulation
2. Structured light
3. Time of Flight (ToF)
4. Stereo vision

Triangulation

Remember trigonometry? If you didn’t find it inherently interesting then, 3D imaging provides a cool application for trig. Namely, by knowing the angles between a point source of light and a sensor receiving the light reflected from the target being measured, one can determine the height of the object at the point of reflection.

Now up the game slightly by projecting not just a point source of light but a laser line, and capture the successive point reflections along that line at whatever resolution is attainable. Thereby one has a mapping in one dimension of the surface for a “slice” of the target.

To achieve the third dimension, in this case the Y dimension, move the target at a defined rate of speed, such as on a regulated conveyor belt or stepping stage, and run another cycle of projection and measured reflections. Repeat until the target’s surface has been fully mapped into a point cloud.

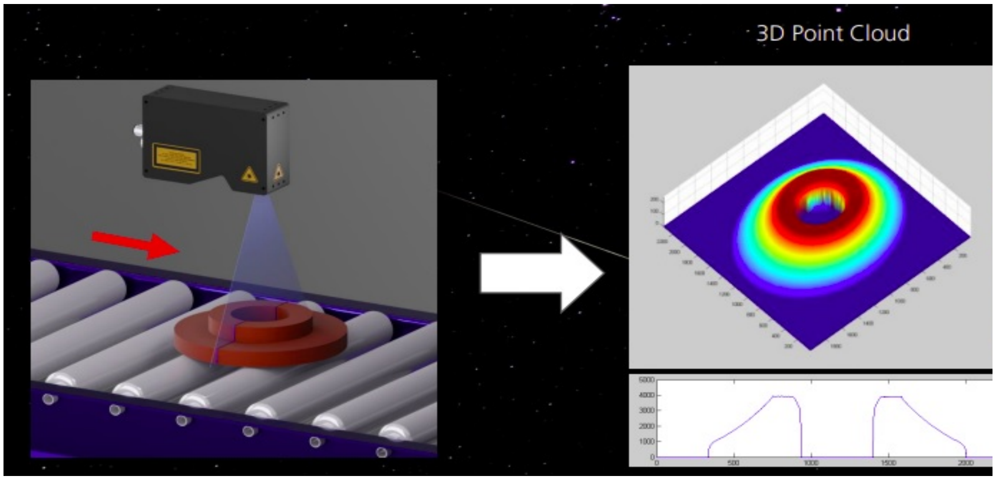

Figure 1: Point cloud generated by successive scans - image courtesy of Automation Technology

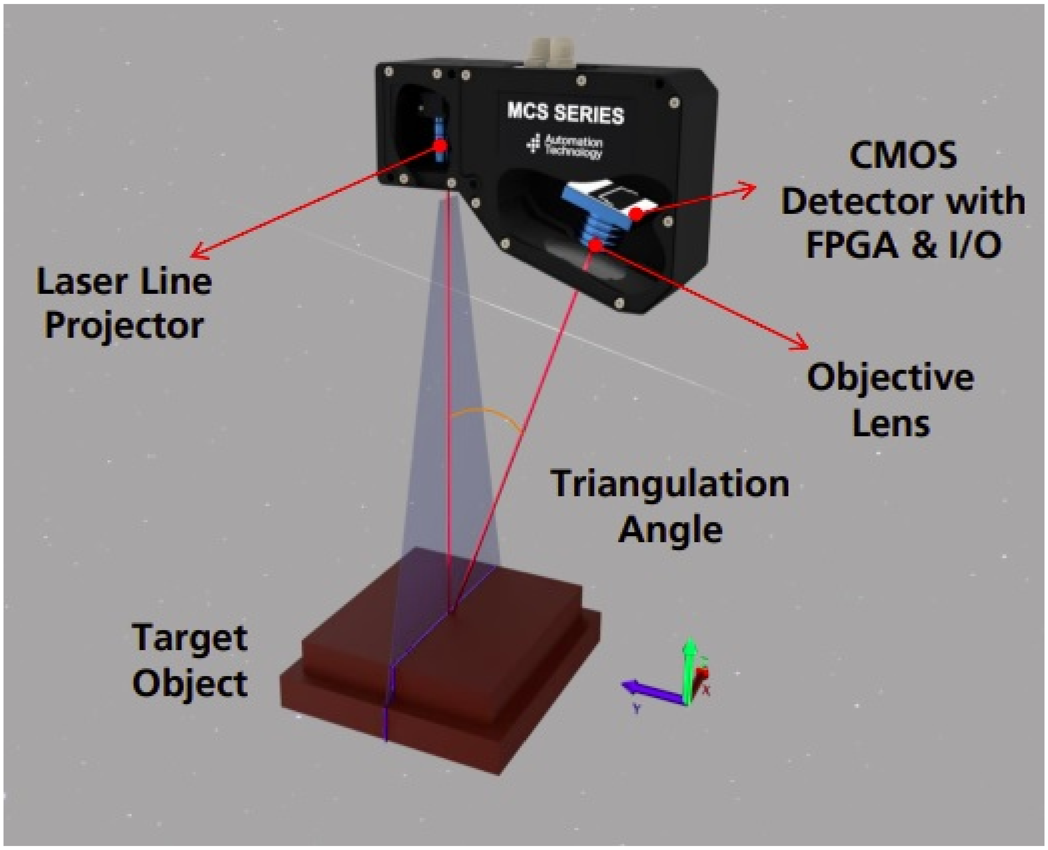

Figure 2: Image courtesy of Automation Technology

So which applications are suited to triangulation? We prefer to complete the survey of all four methods, so that the pros and cons can be presented in full context, by way of comparison. Jump straight to "Comparing the four methods" if you prefer to read the piece by starting with the conclusions.

Structured Light

The structured light methods for 3D imaging follow logically on the triangulation approach surveyed above, with the following key differences. Triangulation creates the Y dimension by moving the target, with the imager mapping the X and Z dimensions.

Structured light applications typically hold the target static, projecting dense patterns, often dots, stripes, or grids, onto the (X,Y) dimensions of the target. In Figure 3 below, dot patterns are projected from a left and a right projector, respectively, with precisely defined offset and projection angles. Sensing how the projected patterns reflect back relative to each other yields the Z data.

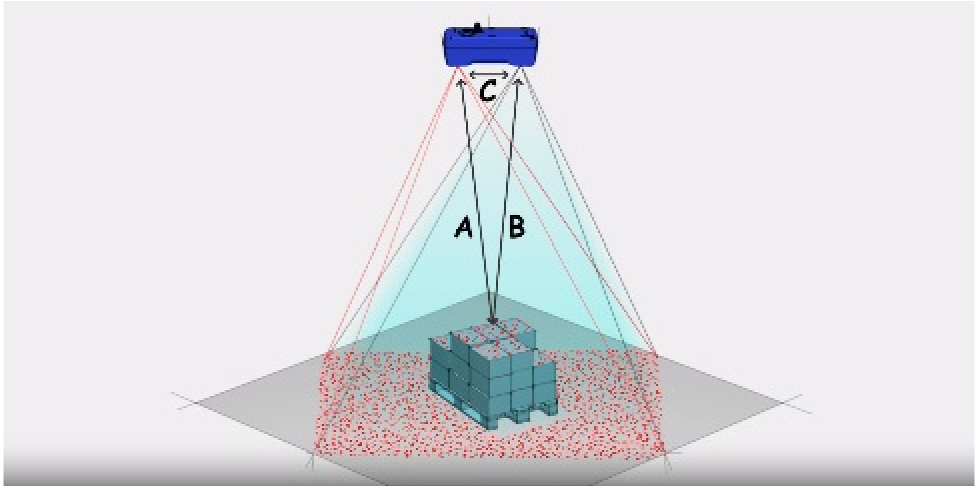

Figure 3: Image courtesy of IDS Imaging

The astute reader will have figured out that the structured light approach also makes use of trigonometric methods like triangulation, so for completeness one might think of the triangulation method previously discussed as "simple triangulation".

Time of Flight (ToF)

If you've got a basic concept for the principles of RADAR (RAdio Detection And Ranging) or SONAR (SOund Navigation And Ranging), you'll find Time of Flight (ToF) imaging easy to understand. Speaking of those analogies, LIDAR means LIght Detection And Ranging and is the umbrella concept for Time of Flight imaging, which is the more commonly used term.

In particular, if one knows the speed at which a signal passes through a medium, by measuring the time it takes for a reflected signal to return to a sensor at the position of the transmitter, one can gauge the distance of the target.

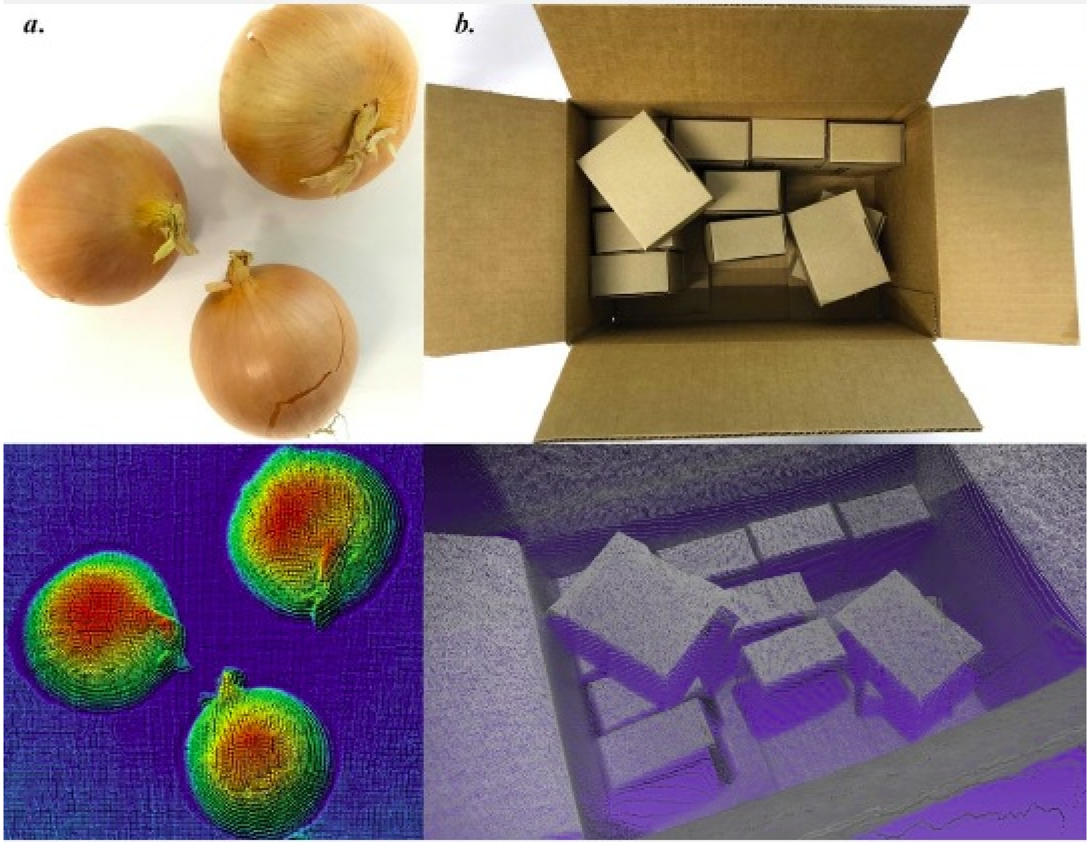

Figure 4: Image courtesy of Lucid Vision Labs

In Figure 3 above:

- Column (a) shows visible light 2D image of onions – top; and 3D point cloud representation of the onions – bottom
- Column (b) shows visible light 2D image of boxes within a larger carton – top; and 3D point cloud representation - bottom

Many Time of Flight systems use light in the near infrared (NIR) range, about 850nm, since it is effective, and has the advantage that it's invisible to the human eye, causes no annoyance when humans are around it, and is safe to humans. As an added bonus, a sensor can be bundled into the same unit to capture a 2D image at the same time, helpful for constructing composite images, QA logging, or other purposes.

Stereo Vision

Just as we two-eyed humans develop depth perception via binocular (aka stereoscopic) vision, each 2D camera sensor obtains an image of the same scene – but from a different angle. For stereo imaging, algorithmically the scenes are correlated and the 3rd dimension is calculated based on the difference in lateral position between the two respective pixels representing the same target feature.

Figure 5: Image courtesy of IDS Imaging

Comparing the four methods:
Which is best? There's no general answer to that. Ultimately it comes down to which is best for your particular application, relative to performance requirements, immutable constraints, cost to implement, competitive benefits of the application in the marketplace, etc.

That said, there are general characteristics associated with each of the four methods. The values cited below are representative at the time of writing (Fall 2023) – we'll update periodically if necessary, as things evolve.

Z-dimension resolution of target: (we do not address (X,Y) resolution here, as it's beyond the scope of this piece; see our other resources on 2D resolution, sensors, and optics, or call us to discuss your overall requirements)

- Triangulation:
    - Z: as fine as 0.2um
- Structured light:
    - Z: As precise at 2.4mm
- Time of Flight:
    - Z: no better than about < 1mm
- Stereo:
    - Z: 0.1mm at 1m; 0.2mm at 2m

Distance of camera(s) from target, (WD):

- Triangulation: 25 - 650mm
- Structured light: 0.5 - 3m
- ToF: 0.3 – 8.3m
- Stereo: 0.5 - 5m

Cost: (relative to each other, but with a breadth of products in each category, and pricing is competitive within and across categories, so let ROI for your application drive the analysis)

- Triangulation: Usually towards the lower cost end of 3D imaging
- Structured light: Usually above median 3D cost point due to system complexity
- ToF: Typically the lowest cost
- Stereo: Usually above median 3D cost point due to system complexity

Bottom line? It comes down to resolution, feature set, and configuration (flexibility), on the performance side, and component and manufacturing cost, from the producer. Pricing ranges from $1500 USD - $17K USD, at single-system pricing.

Other considerations:

- Stereo is well-suited for non-textured surfaces
- ToF not ideal for specular (highly reflective) surfaces, since the entire scene is imaged at once, and multiple reflection paths can result in the measured distance being greater than the actual distance
- Structured light is also best avoided for highly reflective surfaces
- Simple triangulation method only makes sense when target can be moving, as it's the profiles build from successive laser line projections that build up the 3D surface model

---

Summary:

Whether 1D, 2D, or 3D, first determine the application for which you need a solution, and into which of those three categories the solution lies. Look to our other resources, or just call us, for 1D and 2D solutions. For 3D, note the key characteristics of your application, in particular:

- Is the target stationary or moving?
- What are the dimensions of the target?
- How close to the target may the camera get, or how far away could it be?
- What level of measurement precision is required?
- What is an acceptable time frame per cycle?
- Is the project budget driven by performance and the competition, or by cost-control, or a blend of the two in what proportions?
- Do we need a turnkey system, an integrator, or will we build our own system from components?
- In the "Comparing the four methods" section above, does a particular method look most promising?

1stVision represents several camera manufacturers who in turn have various sensors in cameras designed for different applications. We can guide our customers to a solution matched to the machine vision application requirements.

Note: This Tech Brief provides an overview of 3D imaging techniques. There is much more information 1stVision can share with you to be sure you are taking all aspects of the sensor and camera selections into consideration – indeed your entire vision application. We have additional resources we can share to help, so don't hesitate to contact us for free consultation.

1st Vision's sales engineers have over 100 years of combined experience to assist in your camera selection. With a large portfolio of lenses, cables, NIC cards and industrial computers, we can provide a full vision solution.

1stVision, Inc.
info@1stVision.com
40 Shattuck Road, Suite 227
Andover, MA 01810
Office 978-474-0044
Fax 978-268-5743
West Coast 949-361-0350
Canada 519-963-4800

Revised: 2023-September-26