



3D VISION IN INDUSTRIAL ROBOTS

S. Pandiyarajan^{a*}, M. Mugilvalavan^b, M. Maharajan^c

^{a,b}Department of Mechatronics Engineering, Agni College of Technology, Chennai 600130, Tamilnadu, India

^cDepartment of ECE, Agni College of Technology, Chennai 600130, Tamilnadu, India

^a e-mail: pandiyarajan.mht@act.edu.in, ^b e-mail: mugilvalavan.mht@act.edu.in

ABSTRACT

With digital computers development possibility for change of certain human intellect based action raised. Because of a need for human vision replacement, needs that require simplicity, speed and low price in robotized mounting systems, usage in objects recognition and their settlement on exact place robot vision systems took place. This paper explains 3D vision systems implementation industrial robotics. Here are given elements of robot vision systems, ways of assembly and implementation examples.

INTRODUCTION

Industrial robots have existed for several decades, finding an initial foothold in the automotive sector. With tasks requiring high levels of repeatability, large payloads and moderate speeds, automotive plants were ideal for early industrial robots. Today, robots are being deployed in increasing numbers into other industries, including food, pharmaceutical and aerospace. The International Federation of Robotics (IFR) predicts that more than 207,000 industrial robots will be installed worldwide during 2015. Many industrial robots are now being equipped with 3D vision systems. Emerging applications not only require increased speed from industrial robots, but also that the robot be able to locate parts positioned randomly on moving conveyors or stacked in bins or on pallets. Machine vision systems, which have also existed for decades are being paired with robots to help automation systems process such parts. Vision guided robotics (VGR) is fast becoming an enabling technology for the automation of many different processes within many different industries. The technology can be broken down into two main subsets: 2D and 3D. A 2D VGR system will process randomly located parts in a flat plane relative to the robot. A 3D VGR system will process parts randomly located across three dimensions (i.e., X-Y-Z) and is also able to accurately discover each part's 3D orientation. In practice, 2D machine vision is usually performed with a digital camera and software that analyzes a digital image of the part's 2D location and orientation for robotic handling or processing (as in welding or dispensing). There are several different ways to implement 3D vision applications. Moreover, the diversity of existing sensor technologies spans a wide spectrum of resolution and cost.

STEREO VISION

The 3D information for a single part can be obtained by observing a common feature from two different viewpoints. A line in space can be calculated from each of the respective views of the feature being analyzed. The intersection of these two lines at the part feature will yield the respective X-Y-Z values. If multiple features are located on the same part, the part's complete 3D orientation information can be calculated. For example, Figures 1-3 show how two fixed cameras locate an X-Y-

Z point on a part by calculating the intersection of two lines in 3D space from each camera to the part's common feature.

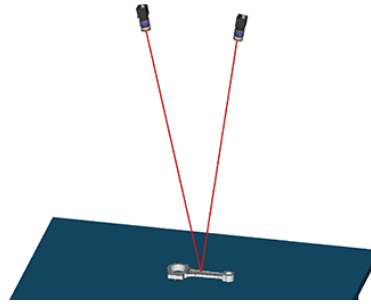


Fig. 1. Stereo vision with two fixed cameras

The cost of a 3D stereo vision system for VGR can be very inexpensive; the system can be built with a single 2D camera if the camera is mounted on a robot, which can then move the camera to two different points of view. The Z-resolution for stereo imaging is dependent on the distance between the two cameras' independent viewpoints. The main disadvantage, though, is that only a single part can be located per "snap."

Stereo vision using structured light

By adding structured light to a stereo vision system, features of the part are not located by identifying features via multiple cameras, but rather through each camera locating the object's features created by the structured light. Typically, this is performed over a large area, with many 3D points detected for each snap. The result is a point cloud, which compiles several X-Y-Z locations in a tight array across the 3D scene being imaged. 3D vision systems that generate point clouds are very useful for VGR applications, because multiple parts can be located simultaneously. The cost of structured-light stereo vision systems varies widely depending on the resolution of the resultant point cloud and the standoff required for the sensor. Examples of structured-light stereo vision systems include Fanuc's 3D Area Sensor, Ensenso's N10 and N20 series of stereo cameras, and Microsoft's Kinect sensor (version 1.0). All of these sensors generate a point cloud based on structured light and stereoscopic vision. With some sensors, the structured light is composed of random dots of IR light. With other sensors, the structured light is very ordered and uses visible light. Figure 4 shows an example of a point cloud.

LASER PROFILER

Another approach to generating 3D data is by analyzing how a laser line falls on a part or on several parts. The laser profile is measured using a 2D vision sensor; given the distance between the laser and the sensor, 3D information can be obtained. To obtain the full 3D information for a significantly sized volume, either the laser profile needs to move across the scene, or the laser itself needs to be moved or pivoted (as is the case with Sick AG's PLB sensor), and then scanned. This way, multiple 3D line scan images can be combined to yield a point cloud of the observed scene. Figure illustrates the PLB sensor being used for robotic bin picking; Figure 6 shows the 3D Area Sensor.

Time of flight

Time of flight (TOF) 3D sensors measure the time it takes for light to travel to the scene and back to individual sensors within a sensor array, similar to the way pixels in a CCD or CMOS vision sensor function. By this method, Z-axis information is obtained from each sensor within the array, and a point cloud is constructed. Different techniques are used to measure the TOF, as measuring the time it takes light to travel such small distances is not possible with current technology. Typically, the phase shift in the emitted light versus the received light provides enough information to calculate the time difference; distance can be calculated by applying this information against the speed of light's known value. Although TOF sensors provide a lower Z-resolution (approximately 1 to 2 cm) than other 3D

sensors, the frame rates possible with TOF technology are much higher than what is possible with the other sensors.

Analyzing 3D data

Once the data is obtained from any 3D sensor, it needs to be analyzed to provide location and orientation information that will make sense to an industrial robot. There are many open-source point cloud libraries readily available for analyzing and extracting useful information from a 3D point cloud to guide an industrial robot. For example, MVTec Software GmbH offers software with 3D analysis tools and APIs that work with a variety of point cloud formats and software languages to make sense of 3D data. Fanuc has a 3D structure stereo sensor that is set up and programmed via the company's iRVision solution. Within it, there are dedicated 3D area sensor vision processes that allow an engineer to quickly calibrate and locate parts for bin picking, depalletization or general 3D guidance. The aforementioned PLB sensor offers higher-level tools with the parameter adjustments required to make sense of the 3D data for robotic guidance.

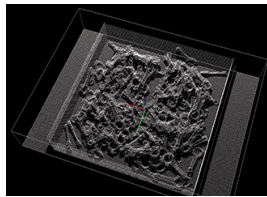


Fig.2. Point cloud generated by Fanuc's 3D Area Sensor

When locating a part's position within a point cloud, a part's 3D CAD data can be used as a model to search for it within the point cloud. It's a CPU-intensive process requiring very advanced algorithms and a lot of computing power. Another approach is to look only for continuous flat surfaces, sometimes referred to as 3D "blobs."

For example, 2D blobs work by looking for 2D pixels, which are connected and similar in some way, with a 2D image that is typically grayscale. With a 3D point cloud, the pixel similarity is the distance from a flat plane. Such algorithm types can be executed much more quickly than full 3D model matching. When an area blob is located, the part's exact location and orientation are not necessarily known, but often this information is enough to guide the robot's tooling to an area of the part that can be gripped in order to remove it from the bin or container.

When sending an industrial robot into a container-like bin, it's important to check the target position and orientation of the robot against the physical constraints of the bin walls. Depending on how a part is located and oriented in a bin, and how the robot and its tooling are supposed to engage the part to extract it from the bin, the robot may not be able to pick the part without colliding with the bin. Some integrated bin-picking software packages consider different "pick positions" for any particular part's location and orientation in order to ensure that the robot will not collide with the bin and, if a collision-free solution cannot be found, that the part can be ignored and the next discovered part can be evaluated. If the robot determines that no parts are able to be picked, a new 3D scan is executed, and this new set of data is examined. If after two scans, no parts are deemed able to be picked, alternative approaches to retrieving the parts need to be investigated. This often involves repositioning parts in the container rather than picking, which requires a different 3D algorithm to locate a feature of a part or perhaps just its highest point. The system then guides the robot to this point, and a path is constructed to reposition the part, drag it away from the walls or simply knock it down off of the wall. One of the biggest challenges with robotic bin picking is quantifying the operation's cycle time. With most traditional robotic applications, the cycle time will be very repeatable (i.e., the robot's motion and entire cycle time is extremely repeatable if parts are precisely located and there is always a place to put down what has been picked up). In practice, there will always be situations in 3D bin-picking applications when a part is not picked for example, if other parts occlude it, or if the robot collides

with the piece and then must back up to re-evaluate the scene. If additional parts are found, the software will shift to the next part and try again; if there are no more found parts, then the robot will leave the bin area and force a rescan.



Fig. 3. A graphic of Sick's PLB robotic bin picking system.

These situations create cycle-time variability which, depending on the next step in the manufacturing process, may not be well-tolerated. Often, bin-picking systems have a part buffer installed between the bin-picking robot and the rest of the automation process to absorb these periodically higher cycle times. The tooling used to grab a part and remove it from the bin is very specialized to the part type and the application; the specific technology used can dictate success or failure. Vacuum- and magnetic-based grippers are the preferred technology for bin-picking tooling, as they allow the robot to contact the part on a variety of surfaces and do not require the tooling to descend too far into the mass of parts. This reduces the likelihood of colliding with other parts. Some parts will only work with a mechanical grip, though, and there are 3D algorithms written specifically for mechanical grippers. These algorithms look for white space within the point cloud, or volumes of no points where the mechanical gripper fingers will fit within point cloud data existing between these areas. In this approach, the gripper is guided to an area of the target part that can be gripped, then extracts the part, although the part's precise location and orientation are not known. They can, however, be discovered with a secondary 2D or 3D vision process performed by another robot or via the same robot, should the cycle time permit.



Figure 4. 3D Area Sensor for iRVision.

When it comes to 3D VGR, selecting the correct 3D vision technology is important; the type chosen depends on the specific application. Some 3D VGR applications do not require the parts to be accurately picked by the robot, but simply that each part to be removed from its container or pallet and placed randomly on a conveyor or table. For these types of applications, TOF 3D sensors make sense, especially if the required throughput is high. For applications where the part needs to be loaded very accurately to the next stage in the automation process. It's possible, though, to use less-accurate 3D information to remove a part from a bin and then use a secondary 2D vision system to achieve a more-precise part location. This two-step process is often used for 3D VGR bin-picking applications, typically due to the fact that even if the part can be accurately located in the bin, it may not be possible for it to be picked correctly because the part's preferred pick location, as determined by the robot, is inaccessible to the system. Sensors measuring higher cycle times (i.e., low frame rates) are well-suited for bin-picking applications because they enable many parts to be located from one scan. While the robot is processing the part it has just picked from the container, there is ample time for the sensor to scan the bin for more parts. Here, Z-resolution information is important, but often errors can be

compensated for within the tooling design. Typically, there is some travel built into the gripper in the direction of the Z axis, which allows for the robot to absorb the error without causing a collision. In a vacuum tooling system, the actual vacuum cups have bellows that act as Z-compensators and, as the robot attempts to pick a part, the vacuum switch can be continuously monitored so that as soon as a sufficient vacuum is generated, the robot will stop and then start its retreat path out of the bin. Additionally, a robot can sense when it comes in contact with an object without necessarily setting off a collision alarm, provided that robot's speed is low (i.e., <100 mm/s). This is done by carefully monitoring the torque on each robot's axis for values that exceed expected values. Cost of the sensors can be an issue depending on the application, as can be the weight of the part and the type of technology being used to remove the parts. Applications involving heavy parts being processed for two to three shifts per day can tolerate a higher-cost sensor, as the return on investment for automating such a task will be quick.

REFERENCES

- [1] V. Doleček and I. Karabegović, "Robotika", *Tehnički fakultet Bihać*, 2002.
- [2] Z. Kovačić and K. Petrinc, "Robotski vid", *Fakultet Elektrotehnike i računarstva Zagreb*, 2005.
- [3] S. Wienand, "Robot vision for Depalletizing of Mixed Pallet Loads Isra Vision Systems", 2006.
- [4] B. Jerbić, B. Vranješ, M. Hrman and Z. Kunica, "Inteligentno robotsko sklapanje primjenom vizijskog sustava i CAD integracije", *Transactions of Famena XXIX-1 Fakultet strojarstva i brodogradnje Zagreb*, 2005.
- [5] D. Z. Miljković, "Sistemi veštačkih neuronskih mreža u proizvodnim tehnologijama", *Mašinski fakultet Univerziteta u Beogradu*, 2002.
- [6] S. Lončarić, "Digitalna analiza slike", *Fakultet Elektrotehnike i računarstva Zagreb*, 2002.
- [7] Available: <http://www.isravisio.com>.
- [8] D. Vernon, "Machine Vision Automated Visual Inspection and Robot Vision" in , New York, London:Prentice Hall, 1991.