From MOTION SYSTEM DESIGN (DECEMBER 2007)

Visual Servoing – Closing the Position Loop with Vision

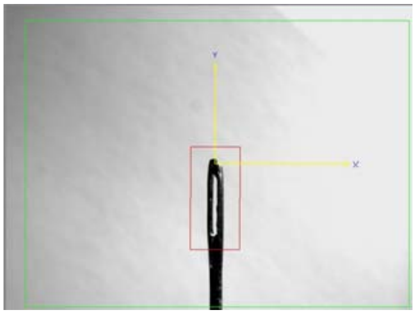
Jim Shimano, Product Manager
Precise Automation, Los Altos, CA

As products shrink in size, the cost of automated systems that can satisfy the requirements is skyrocketing. By using machine vision to close a robot's position loop, high accuracy automated systems can be created using low-cost, high-resolution, low accuracy robots. This process, known as visual servoing, lowers cost, increases flexibility and automates processes that could not be cost effectively automated before.

Products in many industries are shrinking to the point where people can no longer reliably handle or assemble them. This trend is accelerating and is creating a demand for automated systems that can feed, align and fasten together complex, 3D electro-mechanical-optical parts. A key component in these systems is vision guidance, using machine vision to locate parts that are to be accessed by a robot. Consumer demand for digital and cell phone camera technology has greatly reduced the cost of high-resolution industrial cameras to the point where they can be cost effectively integrated into low-cost automated systems. These cameras, combined with improved processing speeds and more powerful vision algorithms, make machine vision an increasingly effective way to locate and to identify parts regardless of their position, orientation or size.

While traditional vision guidance is an excellent way to locate parts, its benefit in high precision assembly operations is limited by the accuracy of the robot it controls. For applications such as the placement of a 1 mm^2 laser diode or a semiconductor disk drive read head, the accuracy requirements can be as small as one or two microns. To achieve placement accuracies in this range, an extremely expensive, high accuracy robot must be used. While many robots have high-resolution encoders, their ability to move to an absolute XYZ position commanded by a traditional vision system can be quite limited due to manufacturing variations, thermal expansion and a host of mechanical effects.

However, with visual servoing, the process of visually closing the robot's position loop, a robot can achieve placement accuracies based on its encoder resolution rather than its absolute accuracy. Now applications such as placing laser diodes into DVD read heads, locating laser diodes relative to an optical lens or teaching wafer slot positions in a semiconductor wafer carrier can be effectively automated using available, low-cost, high-resolution robots that have limited intrinsic absolute accuracy.

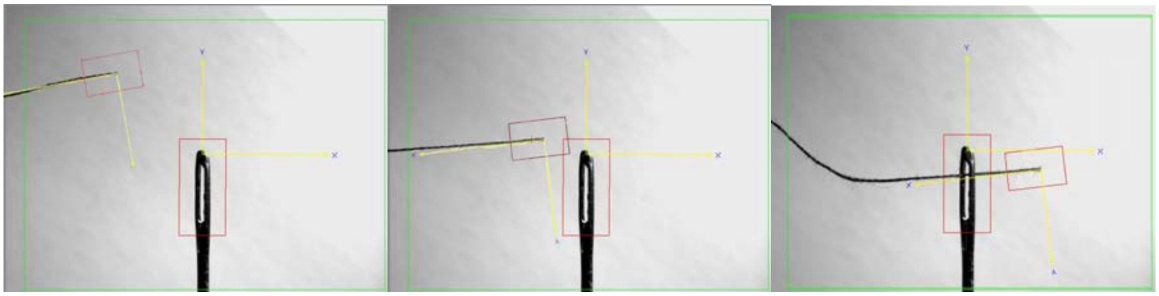


Traditional vision guidance – system captures a single image of the target and transmits a position and orientation in world coordinates.

Threading the needle

In a traditional vision guidance application, the vision system captures a single picture and locates a target's position in world coordinates. Vision transmits these coordinates to a robot and relies on the mechanism to move to the position accurately. This is analogous to a person looking at a part and then grabbing it with their eyes closed. If the part is big enough, this may be successful. However, for many applications, such as threading a needle, this process could not be used. This is analogous to robotic applications with small parts.

So how can vision be used to thread a needle? A person carries out this task by moving the thread towards the eye of the needle. By looking at both objects simultaneously, they determine their relative distance and direction and move the thread accordingly. Robotic visual servoing works in the same manner. The vision system takes a picture and analyzes the relative positions of needle and thread. Instead of sending the robot a single motion command in world coordinates, the software sends a series of incremental distance and direction motion commands. As the motion is executed, more pictures are taken. The software continuously analyses the new images and updates the motion commands accordingly. This process continues until the vision system confirms that the task is accomplished.



Visual Servoing – system captures multiple images and determines the relative distance between the part and the target. A series of incremental motion commands are executed until vision confirms the task is accomplished.

Resolution, not accuracy

In a traditional vision guidance system, the software takes a single picture and provides the motion command to the robot in world coordinates. It is the robot's responsibility to move to this location accurately and repeatably. The success of this process relies upon many factors:

- An accurate knowledge of the position of the camera relative to the robot.
- An accurate knowledge of where the camera is looking (its field of view).
- The robot's ability to accurately move to a commanded XYZ position after taking into account the effects of link thermal expansion, drive train wear, manufacturing tolerances in the straightness and perpendicularity of its links, run out and backlash, etc.

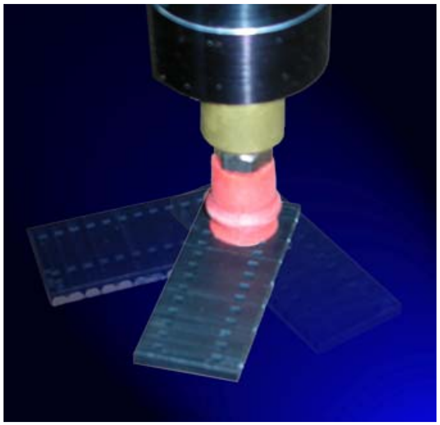
Since there is no verification that the motion is executed correctly, the system must assume the process is successful and progresses to the next step. If a traditional vision guidance application involves very small parts, accurate and expensive robots are needed to reliably automate the process. However, even the most accurate mechanisms still require periodic recalibration of the entire system.

Visual servoing replaces a single world coordinate command in favor of a series of distance and direction commands, so the need for absolute accuracy is reduced. The robot executes a motion by traveling a specified incremental distance. As the motion progresses, the system takes more pictures, updating the motion command with every image. Since the motion will continue until the vision system confirms that the target has been reached, it is not important if the robot is not able to exactly move to each requested position. The visual servoing loop will adjust for any differences. Thus, as long as the resolution of the camera capturing the image is as good as or better than the resolution of the robot's encoders, the system can achieve placement accuracies equal to the encoder resolution. Visual servoing also confirms that motions are properly executed before advancing to the next step.

End effecters made easy

Since traditional vision-guidance systems only observe the target location and not the part as it is captured by the robot, the system must repeatedly capture parts in an identical manner every cycle. This is rarely achieved if the system uses low-cost end effectors, such as vacuum grippers. Thus, traditional systems must rely on expensive and often times custom-made grippers to capture parts reliably.

A visual servoing system can adapt to inaccurate grippers. Since both the part and target are viewed and since only their relative positions are important, the part can be gripped in a range of positions or orientations. This flexibility can also be used to automate processes where parts and targets change in real time, such as the placement of a read head on a disk drive suspension or the insertion of sutures into the hole of suture needles. As the part or target change, the vision system can detect the variations and adjust accordingly.



Visual servoing will automatically compensate for inaccurate grippers such as vacuum end-effectors. The part can be gripped in a range of positions or orientations and the vision system will adjust.

Variation in camera position and orientation

Traditional vision guidance requires an accurate calibration between its field of view and the robot's coordinates. If the camera is fixed mounted on a stand, the camera stand is susceptible to the same thermal expansion and age inaccuracies as a robotic mechanism. Even if the camera is mounted on the robot, for traditional vision guidance, an accurate calibration is still required and it may vary over time.

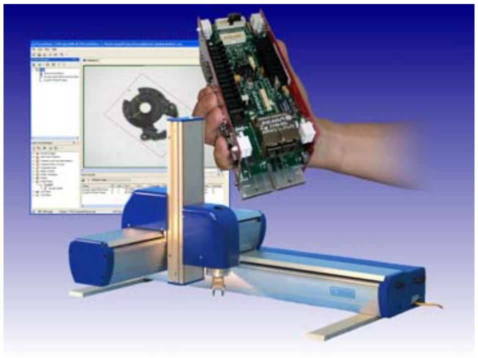
Since visual servoing favors relative distance over world coordinates, the needed calibration is simpler. As long as the relationship between pixels and real distance stays relatively constant, small disturbances in the camera mounting do not affect the ultimate accuracy and performance of the system. Since visual servoing is effectively a dynamically adaptive calibration, the need to recalibrate the robot-vision system is greatly reduced.

Lens and lighting

Visual servoing requires that both the target location and the part in the end effecter are viewed simultaneously. Since these two objects are often in different Z planes, the choices of lens and lighting used are very important. A lens with a large depth of field is useful, but it is necessary to choose a focal distance that either fully views the part, the target or an area between the two. In any case, one or both of the objects may be slightly out of focus making it important to have a powerful vision software package that features an extensive toolkit.

The importance of cycle time

For visual servoing to be effective, the cycle time from image capture to robot motion is critical. Since the system must process multiple images for every assembly operation, it is important to have a system that can successfully complete the image-processing to robot-motion cycle in less than 100 milliseconds. Faster processors and improved vision algorithms make these speeds possible. However, a visual servoing solution created with mix and match components will typically suffer from vision-system to motion-system communication delays of 50 milliseconds. These communication delays added to processing and motion time result in only five to ten frames/second throughput. To take advantage of visual servoing, a tightly integrated robot, controller and vision software system is required. With a fully integrated system, communication delays have been reduced to as little as two milliseconds. By using efficient messaging, as many as 17 frames/second can be processed. Thus, an integrated robot, controller and vision package insures that you can take full advantage of the lower cost and increased flexibility that visual servoing has to offer.



The cycle time from image capture to robot motion is critical. A tightly integrated robot, controller and vision software system is necessary to take full advantage of visual servoing.