Robots learn to handle objects, understand places

September 2 2011, By Bill Steele

After scanning a room, a robot points to the keyboard it was asked to locate. It uses context to identify objects, such as the fact that a keyboard is usually in front of a monitor.

(PhysOrg.com) -- Infants spend their first few months learning to find their way around and manipulating objects, and they are very flexible about it: Cups can come in different shapes and sizes, but they all have handles. So do pitchers, so we pick them up the same way.

Similarly, your [personal robot](#) in the future will need the ability to generalize -- for example, to handle your particular set of dishes and put them in your particular dishwasher.

In Cornell's Personal Robotics Laboratory, a team led by Ashutosh Saxena, assistant professor of computer science, is teaching robots to manipulate objects and find their way around in new environments. They

reported two examples of their work at the 2011 Robotics: Science and Systems Conference June 27 at the University of Southern California.

A common thread running through the research is "machine learning" -- programming a computer to observe events and find commonalities. With the right programming, for example, a computer can look at a wide array of cups, find their common characteristics and then be able to identify cups in the future. A similar process can teach a robot to find a cup's handle and grasp it correctly.

Other researchers have gone this far, but Saxena's team has found that placing objects is harder than picking them up, because there are many options. A cup is placed upright on a table, but upside down in a dishwasher, so the robot must be trained to make those decisions.

"We just show the robot some examples and it learns to generalize the placing strategies and applies them to objects that were not seen before," Saxena explained. "It learns about stability and other criteria for good placing for plates and cups, and when it sees a new object -- a bowl -- it applies them."

In early tests they placed a plate, mug, martini glass, bowl, candy cane, disc, spoon and tuning fork on a flat surface, on a hook, in a stemware holder, in a pen holder and on several different dish racks.

Placing dishes in a rack is a challenging task for a robot. It must identify empty spaces and place the plate in the correct upright position.

Surveying its environment with a 3-D camera, the robot randomly tests small volumes of space as suitable locations for placement. For some objects it will test for "caging" -- the presence of vertical supports that would hold an [object](#) upright. It also gives priority to "preferred" locations: A plate goes flat on a table, but upright in a dishwasher.

After training, their robot placed most objects correctly 98 percent of the time when it had seen the objects and environments previously, and 95 percent of the time when working with new objects in a new environment. Performance could be improved, the researchers suggested, by longer training.

But first, the robot has to find the dish rack.

Just as we unconsciously catalog the objects in a room when we walk in, Saxena and colleague Thorsten Joachims, associate professor of [computer science](#), have developed a system that enables a robot to scan a

room and identify its objects. Pictures from the robot's 3-D camera are stitched together to form a 3-D image of an entire room that is then divided into segments, based on discontinuities and distances between objects. The goal is to label each segment.

The researchers trained a robot by giving it 24 office scenes and 28 home scenes in which they had labeled most objects. The computer examines such features as color, texture and what is nearby and decides what characteristics all objects with the same label have in common. In a new environment, it compares each segment of its scan with the objects in its memory and chooses the ones with the best fit.

"The novelty of this work is to learn the contextual relations in 3-D," Saxena said. "For identifying a keyboard it may be easier to locate the monitors first, because the keyboards are found below the monitors."

In tests, the robot correctly identified objects about 83 percent of the time in home scenes and 88 percent in offices. In a final test, it successfully located a keyboard in an unfamiliar room. Again, Saxena said, context gives this robot an advantage. The keyboard only shows up as a few pixels in the image, but the monitor is easily found, and the robot uses that information to locate the keyboard.

Robots still have a long way to go to learn like humans, the researchers admit. "I would be really happy if we could build a robot that would even act like a six-month-old baby," Saxena said.

Provided by Cornell University