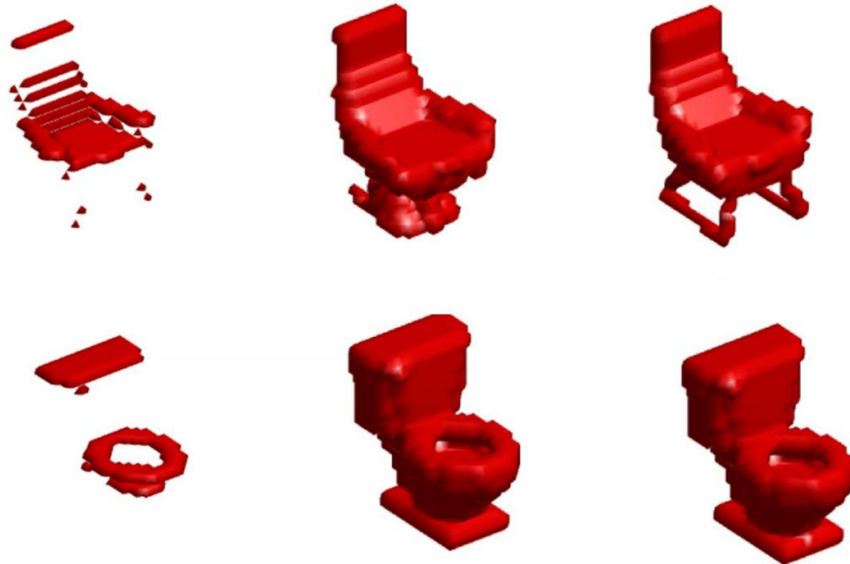# Helping robots learn to see in 3-D

July 14 2017, by Robin A. Smith

When fed 3-D models of household items in bird's-eye view (left), a new algorithm is able to guess what the objects are, and what their overall 3-D shapes should be. This image shows the guess in the center, and the actual 3-D model on the right. Credit: Ben Burchfiel

Autonomous robots can inspect nuclear power plants, clean up oil spills in the ocean, accompany fighter planes into combat and explore the surface of Mars.

Yet for all their talents, robots still can't make a cup of tea.

That's because tasks such as turning the stove on, fetching the kettle and finding the milk and sugar require perceptual abilities that, for most machines, are still a fantasy.

Among them is the ability to make sense of 3-D objects. While it's relatively straightforward for robots to "see" objects with cameras and other sensors, interpreting what they see, from a single glimpse, is more difficult.

Duke University graduate student Ben Burchfiel says the most sophisticated robots in the world can't yet do what most children do automatically, but he and his colleagues may be closer to a solution.

Burchfiel and his thesis advisor George Konidaris, now an assistant professor of computer science at Brown University, have developed new technology that enables machines to make sense of 3-D objects in a richer and more human-like way.

A robot that clears dishes off a table, for example, must be able to adapt to an enormous variety of bowls, platters and plates in different sizes and shapes, left in disarray on a cluttered surface.

Humans can glance at a new [object](#) and intuitively know what it is, whether it is right side up, upside down or sideways, in full view or partially obscured by other objects.

Even when an object is partially hidden, we mentally fill in the parts we can't see.

Their robot perception algorithm can simultaneously guess what a new object is, and how it's oriented, without examining it from multiple

angles first. It can also "imagine" any parts that are out of view.

A robot with this technology wouldn't need to see every side of a teapot, for example, to know that it probably has a handle, a lid and a spout, and whether it is sitting upright or off-kilter on the stove.

The researchers say their approach, which they presented July 12 at the [2017 Robotics: Science and Systems Conference](#) in Cambridge, Massachusetts, makes fewer mistakes and is three times faster than the best current methods.

This is an important step toward robots that function alongside humans in homes and other real-world settings, which are less orderly and predictable than the highly controlled environment of the lab or the factory floor, Burchfiel said.

With their framework, the robot is given a limited number of training examples, and uses them to generalize to new objects.

"It's impractical to assume a robot has a detailed 3-D model of every possible object it might encounter, in advance," Burchfiel said.

The researchers trained their algorithm on a dataset of roughly 4,000 complete 3-D scans of common household objects: an assortment of bathtubs, beds, chairs, desks, dressers, monitors, nightstands, sofas, tables and toilets.

Each 3-D scan was converted into tens of thousands of little cubes, or voxels, stacked on top of each other like LEGO blocks to make them easier to process.

The algorithm learned categories of objects by combing through examples of each one and figuring out how they vary and how they stay

the same, using a version of a technique called probabilistic [principal component analysis](#).

When a [robot](#) spots something new—say, a bunk bed—it doesn't have to sift through its entire mental catalogue for a match. It learns, from prior examples, what characteristics beds tend to have.

Based on that prior knowledge, it has the power to generalize like a person would—to understand that two objects may be different, yet share properties that make them both a particular type of furniture.

To test the approach, the researchers fed the algorithm 908 new 3-D examples of the same 10 kinds of household items, viewed from the top.

From this single vantage point, the algorithm correctly guessed what most objects were, and what their overall 3-D shapes should be, including the concealed parts, about 75 percent of the time—compared with just over 50 percent for the state-of-the-art alternative.

It was also capable of recognizing objects that were rotated in various ways, which the best competing approaches can't do.

While the system is reasonably fast—the whole process takes about a second—it is still a far cry from human vision, Burchfiel said.

For one, both their algorithm and previous methods were easily fooled by objects that, from certain perspectives, look similar in shape. They might see a table from above, and mistake it for a dresser.

"Overall, we make a mistake a little less than 25 percent of the time, and the best alternative makes a mistake almost half the time, so it is a big improvement," Burchfiel said. "But it still isn't ready to move into your house. You don't want it putting a pillow in the dishwasher."

Now the team is working on scaling up their approach to enable robots to distinguish between thousands of types of objects at a time.

"Researchers have been teaching robots to recognize 3-D objects for a while now," Burchfield said. What's new, he explained, is the ability to both recognize something and fill in the blind spots in its field of vision, to reconstruct the parts it can't see.

"That has the potential to be invaluable in a lot of robotic applications," Burchfiel said.

More information: "[Bayesian Eigenobjects: A Unified Framework for 3D Robot Perception](#)," Benjamin Burchfiel and George Konidaris. RSS 2017, July 12-16, 2017, Cambridge, Massachusetts.

Provided by Duke University

Citation: Helping robots learn to see in 3-D (2017, July 14) retrieved 23 April 2024 from https://phys.org/news/2017-07-robots-d.html