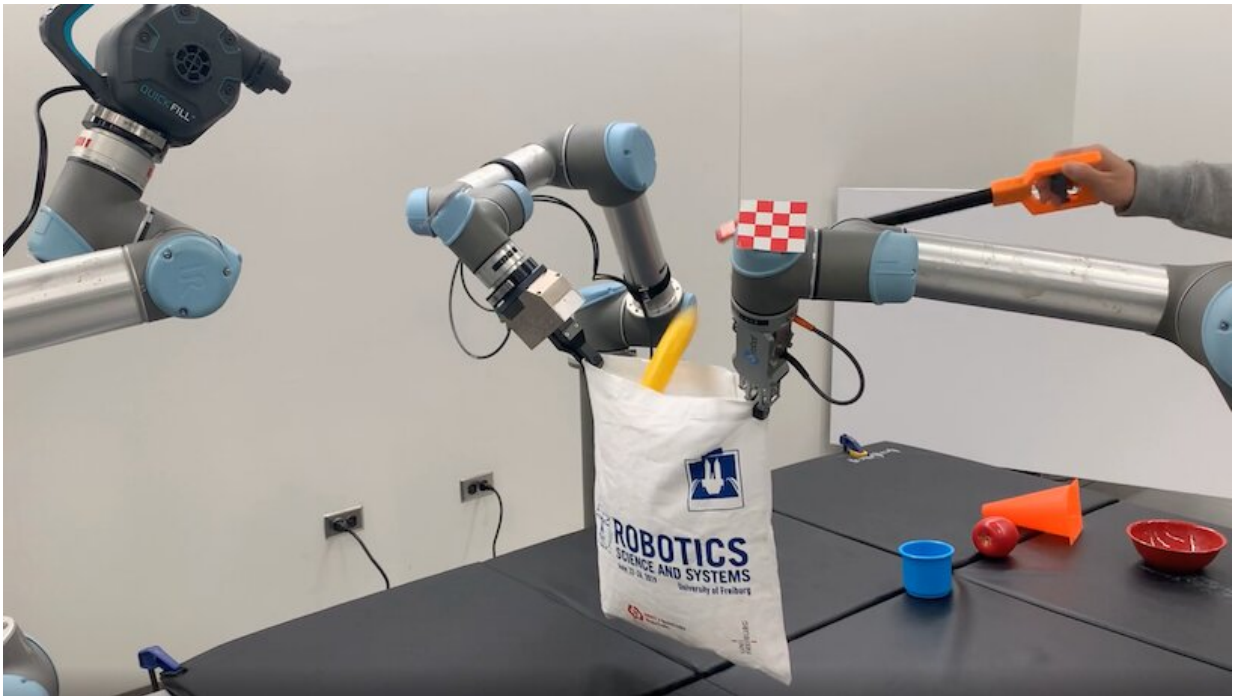


Engineers use psychology, physics, and geometry to make robots more intelligent

March 8 2023, by Bernadette Young



The self-supervised learning framework Columbia Engineers call DextAIRity learns to effectively perform a target task through a sequence of grasping or air-based blowing actions. Using visual feedback, the system uses a closed-loop formulation that continuously adjusts its blowing direction. Credit: Zhenjia Xu/Columbia Engineering

Robots are all around us, from drones filming videos in the sky to serving food in restaurants and diffusing bombs in emergencies. Slowly

but surely, robots are improving the quality of human life by augmenting our abilities, freeing up time, and enhancing our personal safety and well-being. While existing robots are becoming more proficient with simple tasks, handling more complex requests will require more development in both mobility and intelligence.

Columbia Engineering and Toyota Research Institute computer scientists are delving into psychology, physics, and geometry to create algorithms so that robots can adapt to their surroundings and learn how to do things independently. This work is vital to enabling robots to address new challenges stemming from an aging society and provide better support, especially for seniors and people with disabilities.

Teaching robots occlusion and object permanence

A longstanding challenge in computer vision is object permanence, a well-known concept in psychology that involves understanding that the existence of an object is separate from whether it is visible at any moment. It is fundamental for robots to understand our ever-changing, dynamic world. But most applications in computer vision ignore occlusions entirely and tend to lose track of objects that become temporarily hidden from view.

"Some of the hardest problems for [artificial intelligence](#) are the easiest for humans," said Carl Vondrick, associate professor of computer science and a recipient of the Toyota Research Institute Young Faculty award. Think about how toddlers play peek-a-boo and learn that their parent does not disappear when they cover their faces. Computers, on the other hand, lose track once something is blocked or hidden from view and cannot process where the object went or recall its location.

To tackle this issue, Vondrick and his team taught [neural networks](#) the basic physical concepts that come naturally to adults and children.

Similar to how a child learns physics by watching events unfold in their surroundings, the team created a machine that watches many videos to learn physical concepts. The key idea is to train the computer to anticipate what the scene would look like in the future. By training the machine to solve this task across many examples, the machine automatically creates an internal model of how objects physically move in typical environments. For example, when a soda can disappears from sight inside the refrigerator, the machine learns to remember it still exists because it appears again once the refrigerator door reopens.

"I have worked with images and videos before, but getting neural networks to work well with 3-D information is surprisingly tricky," said Basile Van Hoorick, a third-year Ph.D. student who worked with Vondrick to [develop the framework that can understand occlusions as they occur](#). Unlike humans, an understanding of the three-dimensionality of our world does not come naturally to computers. The second leap in the project was not only to convert data from cameras into 3-D seamlessly but also to reconstruct the entire configuration of the scene beyond what can be seen.

This work could expand the perception capabilities of home robots widely. In any indoor environment, things become hidden from view all the time. Hence, robots need to interpret their surroundings intelligently. The "soda can inside the refrigerator" situation is one of many examples. Still, it is easy to see how any application that uses vision will benefit if robots can draw upon their memory and object-permanence reasoning skills to keep track of both objects and humans as they move around the house.

Moving beyond the rigid body assumption

Most robots today are programmed with a series of assumptions for them to work. One is the rigid body assumption, which assumes that an

object is solid and doesn't change shape. And that simplifies a lot of things. Roboticists can completely ignore the physics of the object the [robot](#) is interacting with and only have to think about the robot's motion.

The Columbia Artificial Intelligence and Robotics (CAIR) Lab, led by Computer Science Assistant Professor Shuran Song, has been researching robotic movement in a different way. Her research focuses on deformable, non-rigid objects—they fold, bend, and change shape. When working with deformable objects, roboticists can no longer rely on the rigid body assumption, forcing them to think about physics again.

"In our work, we are trying to investigate how humans intuitively do things," said Shuran Song, also a Toyota Research Institute Young Faculty awardee. Instead of trying to account for every possible parameter, her team developed an algorithm that allows the robot to learn from doing, making it more generalizable and lessening the need for massive amounts of training data. It forced the group to rethink how people do an action, like hitting a target with a rope. We usually don't think about the trajectory of the string—instead, we try to hit the object first and adjust our movements until we're successful. "This new perspective was essential to solve this [difficult problem](#) in robotics," Song noted.

Her team won a Best Paper Award at the Robot Science and Systems Conference (RSS 2022) for an algorithm they developed, Iterative Residual Policy (IRP). IRP is a general learning framework for repeatable tasks with complex dynamics, where a single model was trained using inaccurate simulation data. The algorithm can learn from that data and hit many targets with unfamiliar ropes in robotic experiments, reaching sub-inch accuracy and demonstrating its strong generalization capability.

"Previously, to achieve this level of precision, the robot needed to do the

task maybe 100 to 1,000 times," said Cheng Chi, a third-year Ph.D. student who worked with Song to develop IRP. "With our system, we can do it within ten times, which is about the same performance as a person."

The researchers noticed that there were still some limitations with the flinging motion that their robot could make. While the flinging motion is effective, it is limited by the speed of the robot arm, which means it cannot handle large items. Not to mention that it is dangerous to have a fast flinging motion around people.

Song's team took this research a step further and developed a new approach to manipulating them by using actively blown air. They armed their robot with an air pump and it was able to quickly unfold a large piece of cloth or open a plastic bag. The self-supervised learning framework they call [DextAIRity](#) learns to effectively perform a target task through a sequence of grasping or air-based blowing actions. Using visual feedback, the system uses a closed-loop formulation that continuously adjusts its blowing direction.

"One of the interesting strategies the system developed with the bag-opening task is to point the air a little above the [plastic bag](#) to keep the bag open," said Zhenjia Xu, a fourth-year Ph.D. student who works with Song in the CAIR Lab. "We did not annotate or train it in any way; it learned it by itself."

What needs to be done to make robots more useful in our homes?

Currently, robots can successfully maneuver through a structured environment with clearly defined areas and do one task simultaneously. However, a truly useful home robot should have various skills, be able to

work in an unstructured environment, like a living room with toys on the floor, and handle different situations. These robots will also need to know how to identify a task and which subtasks must be done in what order. And then, they will need to know what to do next if they fail at a job and how to adapt to the next steps needed to accomplish their goal.

"The progress that Carl Vondrick and Shuran Song have made with their research contributes directly to Toyota Research Institute's mission," says Dr. Eric Krotkov, advisor to the University Research Program. "TRI's research in robotics and beyond focuses on developing the capabilities and tools to address the socioeconomic challenges of an aging society, labor shortage, and sustainable production. Endowing robots with the capabilities to understand occluded objects and handle deformable objects will enable them to improve the quality of life for all."

Song and Vondrick plan to collaborate to combine their respective expertise in robotics and [computer vision](#) to create robots that assist people in the home. By teaching machines to understand everyday objects in homes, such as clothes, food, and boxes, the technology could enable robots to assist people with mobility disabilities and improve the quality of everyday life for people. By increasing the number of objects and physical concepts that can be learned by robots, the team aims to make these applications possible in the future.

Provided by Columbia University School of Engineering and Applied Science

Citation: Engineers use psychology, physics, and geometry to make robots more intelligent (2023, March 8) retrieved 23 April 2024 from <https://techxplore.com/news/2023-03-psychology-physics-geometry-robots-intelligent.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.