MIT team's picking and placing system is on another level

March 20 2019, by Nancy Cohen

Credit: Computer Science & Artificial Intelligence Laboratory

If you look at this pick and place robot you promptly see why it is a big deal—not so much for dexterity and fine movement, although the robot scores in both, but just because it is so clever.

It is quite obvious from the news stories pouring out of university labs that robotic arms and hands designed for picking and sorting are a frequent topic; ambitious researchers try to score higher for efficient solutions.

As MIT CSAIL put it, "for all the progress we've made with robots, they still barely have the skills of a two-year-old. Factory robots can pick up the same object over and over again, and some can even make some basic distinctions between objects, but they generally have trouble understanding a wide range of object shapes and sizes, or being able to move said objects into different poses or locations."

This week's buzz is all about this robot, with its featured "keypoints" style for achieving a more advanced level of coordination. They have explored a new way to identify and move *entire classes* of objects, representing them as groups of 3-D keypoints.

The Engineer quoted MIT professor Russ Tedrake, senior author of the paper describing their work and up on arXiv. "Robots can pick almost anything up, but if it's an object they haven't seen before, they can't actually put it down in any meaningful way."

The Engineer gave its nod to an approach that sounded like "a type of visual roadmap that allows more nuanced manipulation."

You can see the robot in action in a kPAM preview video, "Precise robot manipulation with never before seen objects." What is kpam? That stands for Keypoint Affordances for Robotic Level Manipulation. The robot gets all the information it needs to pick up, move and place objects.

"Understanding just a little bit more about the object – the location of a few key points – is enough to enable a wide range of useful manipulation

tasks," said MIT professor Russ Tedrake.

A paper describing their work, which is up on arXiv, is titled "kPAM: KeyPoint Affordances for Category-Level Robotic Manipulation," by Lucas Manuelli, Wei Gao, Peter Florence and Russ Tedrake. They are with CSAIL (Computer Science and Artificial Intelligence Laboratory) of the Massachusetts Institute of Technology.

Here is what the paper's authors had to say about how their approach is a step off existing "manipulation pipelines." The latter typically specify the desired configuration as a target 6-DOF pose, which has its limitations. Representing an object "with a parameterized transformation defined on a fixed template cannot capture large intra-category shape variation, and specifying a target pose at a category level can be physically infeasible or fail to accomplish the task."

Knowing the pose and size of a coffee mug relative to some canonical mug is ok, but it is not sufficient to hang it on a rack by its handle. Their approach is using "semantic 3-D keypoints as the object representation." What were the results of their exploration? Their method was able to handle "large intra-category variations without any instance-wise tuning or specification."

The team reported that "Extensive hardware experiments demonstrate our method can reliably accomplish tasks with never-before seen objects in a category, such as placing shoes and mugs with significant shape variation into category level target configurations."

  More information: kPAM: KeyPoint Affordances for Category-Level Robotic Manipulation, arXiv:1903.06684 [cs.RO] arxiv.org/abs/1903.06684